Alzheimer Disease Genetics Consortium (ADGC), (2019). Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A, tau, immunity and lipid processing. *Nature Genetics*, *51*(3), 414-430. https://doi.org/10.1038/s41588-019-0358-2

Meta-analysis of genetic association with diagnosed Alzheimer's disease identifies novel risk loci and implicates Abeta, Tau, immunity and lipid processing

Kunkle BW+*[1], Grenier-Boley B+[2,3,4], Sims R[5], Bis JC[6], Damotte V[2,3,4], Naj AC[7], Boland A[8], Vronskaya M[5], van der Lee SJ[9], Amlie-Wolf A[10], Bellenguez C[2,3,4], Frizatti A[5], Chouraki V[2,11], Martin ER[1,12], Sleegers K[13,14], Badarinarayan N[5], Jakobsdottir J[15], Hamilton-Nelson KL[1], Moreno-Grau S[16], Olaso R[8], Raybould R[5], Chen Y[17], Kuzma AB[10], Hiltunen M[18,19], Morgan T[5], Ahmad S[9], Vardarajan BN[20-22], Epelbaum J[23], Hoffmann P[24,25,26], Boada M[16], Beecham GW[1,12], Garnier JG[8], Harold D[27], Fitzpatrick AL[28,29], Valladares O[10], Moutet ML[8], Gerrish A[5], Smith AV[30,31], Qu L[10], Bacq D[8], Denning N[5], Jian X[32], Zhao Y[10], Zompo MD[33], Fox NC[34], Choi SH[17], Mateo I[35], Hughes JT[36], Adams HH[9], Malamon J[10], Sanchez-Garcia FS[36], Patel Y[37], Brody JA[6], Dombroski B[10], Naranjo MCD[36], Daniilidou M[38], Eiriksdottir G[15], Mukherjee S[39], Wallon D[40,41], Uphill J[42], Aspelund T[15,43], Cantwell LB[10], Garzia F[8], Galimberti D[44], Hofer E[45,46], Butkiewicz M[47], Fin B[8], Scarpini E[44], Sarnowski C[17], Bush W[47], Meslage S[8], Kornhuber J[48], White CC[49], Song Y[47], Barber RC[50], Engelborghs S[51,52], Pichler S[53], Voijnovic D[9], Adams PM[54], Vandenberghe R[55], Mayhaus M[53], Cupples LA[11,17], Albert MS[56], De Deyn PP[51,52], Gu W[53], Himali JJ[16,57], Beekly D[58], Squassina A[33], Hartmann AM[59], Orellana A[16], Blacker D[60,61], Rodriguez-Rodriguez E[35], Lovestone S[62], Garcia ME[63], Doody RS[64], Munoz-Fernadez CM[36], Sussams R[65], Lin H[17], Fairchild TJ[66], Benito YA[36], Holmes C[65], Comic H[9], Frosch MP[67], Thonberg H[68,69], Maier W[70,71], Roschupkin G[9], Ghetti B[72], Giedraitis V[73], Kawalia A[74], Li S[17], Huebinger RM[75], Kilander L[73], Moebus S[76], Hernández I[16], Kamboh MI[77-79], Brundin R[73], Turton J[74], Yang Q[17], Katz MJ[81], Concari L[82,83], Lord J[74], Beiser AS[11,17], Keene CD[84], Helisalmi S[18,19], Kloszewska I[85], Kukull WA[29], Koivisto AM[18,19], Lynch A[86,87], Tarraga L[16], Larson EB[88], Haapasalo A[89], Lawlor B[86,87], Mosley TH[90], Lipton RB[81], Solfrizzi V[91], Gill M[86,87], Longstreth WT Jr[29,92], Montine TJ[84], Frisardi V[91], Ortega-Cubero S[93,94,95], Rivadeneira F[9,96,97], Petersen RC[98], Deramecourt V[99], Salani F[100], Boerwinkle E[101,102], Reiman EM[103-106], Fievet N[2,3,4], Caltagirone C[100], Rotter JI[107], Reisch JS[108], Hanon O[109], Cupidi C[110], Uitterlinden AG[9,96,97], Royall DR[111], Dufouil C[112,113], Maletta RG[110], de Rojas I[16], Sano M[114], Brice A[115,116], Cecchetti R[117], St George-Hyslop P[118,119], Ritchie K[120,121], Tsolaki M[122], Tsuang DW[123,124], Dubois B[125-128], Craig D[129], Wu CK[130], Soininen H[18,19], Avramidou D[122], Albin RL[131-133], Fratiglioni L[18,19], Germanou A[122], Apostolova LG[134-137], Keller L[69], Koutroumani M[122], Arnold SE[138], Panza F[91], Gkatzima O[122], Asthana S[139-141], Hannequin D[40,41], Whitehead P[1], Makrina D[122], Atwood CS[139-141], Caffarra P[82,83], Hampel H[142-145], Quintela I[146], Carracedo Á[146], Lannfelt L[73], Rubinsztein DC[118,147], Barnes LL[148-150], Pasquier F[99], Frölich L[151], Barral S[20-22], McGuinness B[152], Beach TG[153], Johnston J1[51], Becker JT[77,154,155], Passmore P[152], Bigio EH[156,157], Schott JM[34], Bird TD[92,123], Warren JD[34], Boeve BF[98], Lupton MK[37,158], Bowen JD[159], Proitsi P[37,158], Boxer A[160], Powell JF[37,158], Burke JR[161], Kauwe JK[162], Burns JM[163], Mancuso M[164], Buxbaum JD[114,165,166], Bonuccelli U[164], Cairns NJ[167], McQuillin A[168], Cao C[169], Livingston G[168], Carlson CS[140,141], Bass NJ[168], Carlsson CM[170], Hardy J[171], Carney RM[172], Bras J[34,171,173], Carrasquillo MM[174], Guerreiro R[34,171,173], Allen M[174], Chui HC[175], Fisher E[176], Cribbs DH[160], Masullo C[177], Crocco EA[178], DeCarli C[179], Bisceglio G[175], Dick M[180], Ma L[175], Duara R[181], Graff-Radford NR[175], Evans DA[182], Hodges A[183], Faber KM[72], Scherer M[184], Fallon KB[185], Riemenschneider M[53], Fardo DW[186], Heun R[71], Farlow MR[136], Ferris S[187], Leber M[188], Foroud TM[135], Heuser I[189], Galasko DR[190], Giegling I[59], Gearing M[191,192], Hüll M[193], Geschwind DH[194], Gilbert JR[1,12], Morris J[195,196], Green RC[197], Mayo K[195,198,199], Growdon JH[200], Feulner T[53], Hamilton RL[201], Harrell LE[202], Drichel D[203], Honig LS[20], Cushion TD[5,204], Huentelman MJ[103], Hollingworth P[5], Hulette CM[205], Hyman BT[200], Marshall R[5], Jarvik GP[206,207], Meggy A[5], Abner E[208], Menzies G[5], Jin LW[209], Leonenko G[5], Real LM[210], Jun G[211], Baldwin CT[211], Grozeva D[5], Karydas A[159], Russo G[212], Kaye JA[213,214], Kim R[215], Jessen F[70,71,188], Kowall NW[57,216], Vellas B[217], Kramer JH[218], Vardy E[219], LaFerla FM[220], Jöckel KH[76], Lah JJ[221], Dichgans M[222,223], Leverenz JB[224], Mann D[225], Levey AI[221], Pickering-Brown S[225], Lieberman AP[226], Klopp N[227], Lunetta KL[17], Wichmann HE[228,229,230], Lyketsos CG[231], Morgan K[232], Marson DC[202], Brown K[80], Martiniuk F[233], Medway C[80], Mash DC[234], Nöthen MM[24,25], Masliah E[190,235], Hooper NM[225], McCormick WC[39], Daniele A[236], McCurry SM[237], Bayer A[238], McDavid AN[169], Gallacher J[238], McKee AC[57,216], van den Bussche H[184], Mesulam M[239,157], Brayne C[240], Miller BL[241], Riedel-Heller S[242], Miller CA[243], Miller JW[209], Al-Chalabi A[244], Morris JC[167,198], Shaw CE[244,245], Myers AJ[178], Wiltfang J[246], O'Bryant S[50], Olichney JM[179], Alvarez V[247], Parisi JE[248], Singleton AB[249], Paulson HL[131,133], Collinge J[42], Perry W[1], Mead S[42], Peskind E[124], Rosser M[34], Pierce A[250],Ryan N[42], Poon WW[180], Nacmias B[251,252], Potter H[253], Sorbi S[254,255], Quinn JF[184,185], Sacchinelli E[100], Raj A[169], Spalletta G[100], Raskind M[124], Bossù P[100], Reisberg B[187,256], Clarke R[257], Reitz C[21,22,258], Smith AD[259], Ringman JM[260], Warden D[2], Roberson ED[202], Wilcock G[259], Rogaeva E[119], Bruni AC[110], Rosen HJ[160], Gallo M[110], Rosenberg RN[261], Ben-Shlomo Y[262], Sager MA[140], Mecocci P[117], Saykin AJ[135,137], Pastor P[117], Cuccaro ML[12], Vance JM[12], Schneider JA[148,150,263], Schneider LS[175,264], Slifer S[1], Seeley WW[160], Smith AG[169], Sonnen JA[84], Spina S[72], Stern

RA[57], Swerdlow RH[163], M. Tang[10], Tanzi RE[200], Trojanowski JQ[265], Troncoso JC[266], Van Deerlin VM[265], Van Eldik LJ[267], Vinters HV[260,268], Vonsattel JP[269], Weintraub S[270], Welsh-Bohmer KA[161,271], Wilhelmsen KC[272], Williamson J[20], Wingo TS[221], Woltjer RL[273], Wright CB[274], Yu CE[39], Yu L[148,150], Saba Y[275], Alzheimer Disease Genetics Consortium (ADGC), The European Alzheimer's Disease Initiative (EADI), Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium (CHARGE), Genetic and Environmental Risk in AD/Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease Consortium (GERAD/PERADES), Pilotto A[276], Bullido M.J.[94,277,278], Peters O[188,279], Crane PK[39], Bennett DA[148,150], Bosco P[280], Coto E[277], Boccardi V[117], De Jager PL[281], Lleo A[94,282], Warner N[283], Lopez OL[77,79,154], Ingelsson M[73], Deloukas P[284], Cruchaga C[195,196], Graff C[285,286], Gwilliam R[284], Fornage M[32], Goate AM[165,287], Sanchez-Juan P[35], Kehoe PG[288], Amin N[9], Ertekin-Taner N[174,289], Berr C[120,121,290], Debette S[112,113], Love S[288], Launer LJ[63], Younkin SG[174,289], Dartigues JF[290], Corcoran C[291], Ikram MA[9,292,293], Dickson DW[174], Nicolas G[40,41], Campion D[40,41], Tschanz J[291], Schmidt H[275,294], Hakonarson H[295,296], Clarimon J[94,282], Munger R[291], Schmidt R[45], Farrer LA[17,57,211,297,298], Van Broeckhoven C[13,14], O'Donovan MC[5], DeStefano AL[11,17], Jones L[5], Haines JL[47], Deleuze JF[8], Owen MJ[5], Gudnason V[15,31], Mayeux R[20-22], Escott-Price V[5,204], Psaty BM[10,29,299,300], Ramirez A[74,189], Wang LS[10], Ruiz A~[16], van Duijn CM~[9], Holmans PA~[5], Seshadri S~[11,57,301], Williams J~[5], Amouyel P~[2,3,4], Schellenberg GD~[10], Lambert JC~*[2,3,4], Pericak-Vance MA~*[1,12].

1 John P. Hussman Institute for Human Genomics, University of Miami, Miami, Florida, USA.

2 INSERM, U1167, RID-AGE-Risk Factors and Molecular Determinants of Aging-Related Diseases, Lille, France.

3 Institut Pasteur de Lille, Lille, France.

4 University Lille, U1167-Excellence Laboratory LabEx DISTALZ, Lille, France.

5 Institute of Psychological Medicine and Clinical Neurosciences, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, UK.

6 Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, Washington, USA.

7 Department of Biostatistics and Epidemiology/Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA.

8 Centre National de Recherche en Génomique Humaine (CNRGH), Institut de Biologie François Jacob, CEA, Université Paris-Saclay, F-91057, Evry, France.

9 Department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands.

10 Penn Neurodegeneration Genomics Center, Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA.

11 Framingham Heart Study, Framingham, Massachusetts, USA.

12 Dr. John T. Macdonald Foundation, Department of Human Genetics, University of Miami, Miami, Florida, USA.

13 Neurodegenerative Brain Diseases Group, Center for Molecular Neurology, VIB, Antwerp, Belgium.

14 Laboratory for Neurogenetics, Institute Born-Bunge, University of Antwerp, Antwerp, Belgium.

15 Icelandic Heart Association, Kopavogur, Iceland.

16 Research Center and Memory Clinic of Fundació ACE, Institut Català de Neurociències Aplicades-Universitat Internacional de Catalunya, Barcelona, Spain.

17 Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA.

18 Institute of Biomedicine, University of Eastern Finland, Kuopio, Finland.

19 Department of Neurology, Kuopio University Hospital, Kuopio, Finland.

20 Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Neurology, Columbia University, New York, New York, USA.

21 Gertrude H. Sergievsky Center, Columbia University, New York, New York, USA.

22 Department of Neurology, Columbia University, New York, New York, USA.

23 UMR 894, Center for Psychiatry and Neuroscience, INSERM, Université Paris Descartes, Paris, France.

24 Institute of Human Genetics, University of Bonn, Bonn, Germany.

25 Department of Genomics, Life &Brain Center, University of Bonn, Bonn, Germany.

26 Division of Medical Genetics, University Hospital and Department of Biomedicine, University of Basel, Basel, Switzerland.

27 School of Biotechnology, Dublin City University, Dublin, Ireland.

28 Department of Family Medicine, University of Washington, Seattle, Washington, USA.
29 Department of Epidemiology, University of Washington, Seattle, Washington, USA.
30 Department of Biostatistics, University of Michigan, USA.
31 Faculty of Medicine, University of Iceland, Reykjavik, Iceland.
32 Brown Foundation Institute of Molecular Medicine, University of Texas Health Sciences Center at Houston, Houston, Texas, USA.
33 Section of Neuroscience and Clinical Pharmacology, Department of Biomedical Sciences, University of Cagliari, Cagliari, Italy.
34 Dementia Research Centre, Department of Neurodegenerative Disease, UCL Institute of Neurology, London, UK.
35 Neurology Service and CIBERNED, 'Marqués de Valdecilla' University Hospital (University of Cantabria and IDIVAL), Santander, Spain.
36 Department of Immunology, Hospital Universitario Doctor Negrín, Las Palmas de Gran Canaria, Spain.
37 Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London UK
38 Department of Health Sciences, Psychiatry for the Elderly, University of Leicester, Leicester, UK.
39 Department of Medicine, University of Washington, Seattle, Washington, USA.
40 Normandie University, UNIROUEN, Inserm U1245, Rouen, France.
41 Rouen University Hospital, Department of Neurology, Department of Genetics and CNR-MAJ, F 76000, Normandy Center for Genomic and Personalized Medicine, Rouen, France.
42 Department of Neurodegenerative Disease, MRC Prion Unit, UCL Institute of Neurology, London, UK.
43 Centre for Public Health, University of Iceland, Reykjavik, Iceland.
44 Department of Pathophysiology and Transplantation, University of Milan, Fondazione Ca' Granda, IRCCS Ospedale Policlinico, Milan, Italy.
45 Clinical Division of Neurogeriatrics, Department of Neurology, Medical University Graz, Graz, Austria.
46 Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria.
47 Institute for Computational Biology, Department of Population & Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, USA.
48 Department of Psychiatry and Psychotherapy, University of Erlangen-Nuremberg, Erlangen, Germany.
49 Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA.
50 Department of Pharmacology and Neuroscience, University of North Texas Health Science Center, Fort Worth, Texas, USA.
51 Laboratory for Neurochemistry and Behavior, Institute Born-Bunge, University of Antwerp, Antwerp, Belgium.
52 Department of Neurology and Memory Clinic, Hospital Network Antwerp, Antwerp, Belgium.
53 Department of Psychiatry and Psychotherapy, University Hospital, Saarland, Germany.
54 Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, Texas, USA.
55 Laboratory for Cognitive Neurology, Department of Neurology, University Hospital and University of Leuven, Leuven, Belgium.
56 Department of Neurology, Johns Hopkins University, Baltimore, Maryland, USA.
57 Department of Neurology, Boston University School of Medicine, Boston, Massachusetts, USA.
58 National Alzheimer's Coordinating Center, University of Washington, Seattle, Washington, USA.
59 Department of Psychiatry, Martin Luther University Halle-Wittenberg, Halle, Germany.
60 Department of Epidemiology, Harvard School of Public Health, Harvard University, Boston, Massachusetts, USA.
61 Department of Psychiatry, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts, USA.
62 Department of Psychiatry, University of Oxford, Oxford, UK.
63 Laboratory of Epidemiology and Population Sciences, National Institute on Aging, Bethesda, Maryland, USA.
64 Alzheimer's Disease and Memory Disorders Center, Baylor College of Medicine, Houston, Texas, USA.
65 Division of Clinical Neurosciences, School of Medicine, University of Southampton, Southampton, UK.
66 Office of Strategy and Measurement, University of North Texas Health Science Center, Fort Worth, Texas, USA.
67 C.S. Kubik Laboratory for Neuropathology, Massachusetts General Hospital, Charlestown, Massachusetts, USA.
68 Department of Geriatric Medicine, Karolinska University Hospital Huddinge, Stockholm, Sweden.

69 Aging Research Center, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet and Stockholm University, Stockholm, Sweden.
70 German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany.
71 Department of Psychiatry and Psychotherapy, University of Bonn, Bonn, Germany.
72 Department of Pathology and Laboratory Medicine, Indiana University, Indianapolis, Indiana, USA.
73 Department of Public Health and Caring Sciences/Geriatrics, Uppsala University, Uppsala, Sweden.
74 Department for Neurodegenerative Diseases and Geriatric Psychiatry, University Hospital Bonn, Bonn, Germany
75 Department of Surgery, University of Texas Southwestern Medical Center, Dallas, Texas, USA.
76 Institute for Medical Informatics, Biometry and Epidemiology, University Hospital of Essen, University Duisburg-Essen, Essen, Germany.
77 Department of Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.
78 Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.
79 Alzheimer's Disease Research Center, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.
80 Institute of Genetics, Queen's Medical Centre, University of Nottingham, Nottingham, UK.
81 Department of Neurology, Albert Einstein College of Medicine, Bronx, New York, USA.
82 Section of Neuroscience, DIMEC-University of Parma, Parma, Italy.
83 FERB-Alzheimer Center, Gazzaniga (Bergamo), Italy.
84 Department of Pathology, University of Washington, Seattle, Washington, USA.
85 Elderly and Psychiatric Disorders Department, Medical University of Lodz, Lodz, Poland.
86 Mercer's Institute for Research on Aging, St. James Hospital and Trinity College, Dublin, Ireland.
87 James Hospital and Trinity College, Dublin, Ireland.
88 Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA.
89 A.I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, Kuopio 70211, Finland.
90 Departments of Medicine, Geriatrics, Gerontology and Neurology, University of Mississippi Medical Center, Jackson, MS
91 Department of Geriatrics, Center for Aging Brain , University of Bari, Bari, Italy.
92 Department of Neurology, University of Washington, Seattle, Washington, USA.
93 Neurogenetics Laboratory, Division of Neurosciences, Centre for Applied Medical Research, University of Navarra School of Medicine, Pamplona, Spain.
94 Centro de Investigación Biomédica en Red de Enfermedades Neurodegenerativas (CIBERNED), Instituto de Salud Carlos III, Madrid, Spain.
95 Department of Neurology, Complejo Asistencial Universitario de Palencia, Palencia, Spain.
96 Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands.
97 Netherlands Consortium on Health Aging and National Genomics Initiative, Leiden, the Netherlands.
98 Department of Neurology, Mayo Clinic, Rochester, Minnesota, USA.
99 CHU Lille, Memory Center of Lille (Centre Mémoire de Ressources et de Recherche), Lille, France.
100 Experimental Neuropsychobiology Laboratory, IRCCS Santa Lucia Foundation, Department of Clinical and Behavioral Neurology, Rome, Italy.
101 School of Public Health, Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas, USA.
102 Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, USA.
103 Neurogenomics Division, Translational Genomics Research Institute, Phoenix, Arizona, USA.
104 Arizona Alzheimer's Consortium, Phoenix, Arizona, USA.
105 Banner Alzheimer's Institute, Phoenix, Arizona, USA.
106 Department of Psychiatry, University of Arizona, Phoenix, Arizona, USA.

107 Institute for Translational Genomics and Population Sciences, Departments of Pediatrics and Medicine, Los Angeles BioMedical Research Institute at Harbor-UCLA Medical Center, Torrance, California, USA.

108 Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, Texas, USA.

109 University Paris Descartes, EA 4468, AP-HP, Hôpital Broca, Geriatrics Department, Paris, France.

110 Regional Neurogenetic Centre (CRN), ASP Catanzaro, Lamezia Terme, Italy.

111 Departments of Psychiatry, Medicine, Family & Community Medicine, South Texas Veterans Health Administration Geriatric Research Education & Clinical Center (GRECC), UT Health Science Center at San Antonio, San Antonio, Texas, USA.

112 University of Bordeaux, Neuroepidemiology, Bordeaux, France.

113 INSERM, Neuroepidemiology, UMR 897, Bordeaux, France.

114 Department of Psychiatry, Mount Sinai School of Medicine, New York, New York, USA.

115 INSERM U1127, CNRS UMR 7225, Sorbonne Universités, UPMC Université Paris 06, UMRS 1127, Institut du Cerveau et de la Moelle Épinière, Paris, France.

116 AP-HP, Department of Genetics, Pitié-Salpêtrière Hospital, Paris, France.

117 Section of Gerontology and Geriatrics, Department of Medicine, University of Perugia, Perugia, Italy.

118 Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK.

119 Tanz Centre for Research in Neurodegenerative Disease, University of Toronto, Toronto, Ontario, Canada.

120 INSERM U1061, La Colombière Hospital, Montpellier, France.

121 Montpellier University, Montpellier, France

122 3rd Department of Neurology, Medical School, Aristotle University of Thessaloniki, Thessaloniki, Greece.

123 VA Puget Sound Health Care System/GRECC, Seattle, Washington, USA.

124 Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, Washington, USA.

125 Institut de la Mémoire et de la Maladie d'Alzheimer (IM2A) and Institut du Cerveau et de la Moelle Épinière (ICM), Département de Neurologie, Hôpital de la Pitié-Salpêtrière, Paris, France.

126 Institut des Neurosciences Translationnelles de Paris (IHU-A-ICM), Institut du Cerveau et de la Moelle Épinière (ICM), Paris, France.

127 INSERM, CNRS, UMR-S975, Institut du Cerveau et de la Moelle Epinière (ICM), Paris, France.

128 Sorbonne Universités, Université Pierre et Marie Curie, Hôpital de la Pitié-Salpêtrière, AP-HP, Paris, France.

129 Ageing Group, Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, UK.

130 Departments of Neurology, Pharmacology & Neuroscience, Texas Tech University Health Science Center, Lubbock, Texas, USA.

131 Department of Neurology, University of Michigan, Ann Arbor, Michigan, USA.

132 Geriatric Research, Education and Clinical Center (GRECC), VA Ann Arbor Healthcare System (VAAAHS), Ann Arbor, Michigan, USA.

133 Michigan Alzheimer Disease Center, Ann Arbor, Michigan, USA.

134 Indiana Alzheimer's Disease Center, Indiana University School of Medicine, Indianapolis, Indiana, USA.

135 Department of Medical and Molecular Genetics, Indiana University, Indianapolis, Indiana, USA.

136 Department of Neurology, Indiana University, Indianapolis, Indiana, USA.

137 Department of Radiology and Imaging Sciences, Indiana University, Indianapolis, Indiana, USA.

138 Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA.

139 Geriatric Research, Education and Clinical Center (GRECC), University of Wisconsin, Madison, Wisconsin, USA.

140 Department of Medicine, University of Wisconsin, Madison, Wisconsin, USA.

141 Wisconsin Alzheimer's Disease Research Center, Madison, Wisconsin, USA.

142 AXA Research Fund and UPMC Chair, Paris, France.

143 Sorbonne University, GRC n° 21, Alzheimer Precision Medicine (APM), AP-HP, Pitié-Salpêtrière Hospital, Boulevard de l'hôpital, F-75013, Paris, France.
144 Brain & Spine Institute (ICM), INSERM U 1127, CNRS UMR 7225, Boulevard de l'hôpital, F-75013, Paris, France.
145 Institute of Memory and Alzheimer's Disease (IM2A), Department of Neurology, Pitié-Salpêtrière Hospital, AP-HP, Boulevard de l'hôpital, F-75013, Paris, France.
146 Grupo de Medicina Xenomica, Universidade de Santiago de Compostela, Centro Nacional de Genotipado (CeGen-PRB2-ISCIII), CIBERER, 15782 Santiago de Compostela, Spain.
147 UK Dementia Research Institute, University of Cambridge, Cambridge, UK.
148 Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois, USA.
149 Department of Behavioral Sciences, Rush University Medical Center, Chicago, Illinois, USA.
150 Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois, USA.
151 Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Germany.
152 Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queen's University, Belfast, UK.
153 Civin Laboratory for Neuropathology, Banner Sun Health Research Institute, Phoenix, Arizona, USA.
154 Department of Neurology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.
155 Department of Psychology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA.
156 Department of Pathology, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA.
157 Cognitive Neurology and Alzheimer's Disease Center, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA.
158 Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia.
159 Swedish Medical Center, Seattle, Washington, USA.
160 Department of Neurology, University of California, San Francisco, San Francisco, California, USA.
161 Department of Medicine, Duke University, Durham, North Carolina, USA.
162 Departments of Biology, Brigham Young University, Provo, Utah, USA.
163 University of Kansas Alzheimer's Disease Center, University of Kansas Medical Center, Kansas City, Kansas, USA.
164 Department of Experimental and Clinical Medicine, Neurological Institute, University of Pisa, Pisa, Italy.
165 Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York, USA.
166 Department of Neuroscience, Mount Sinai School of Medicine, New York, New York, USA.
167 Department of Pathology and Immunology, Washington University, St. Louis, Missouri, USA.
168 Division of Psychiatry, University College London, UK
169 USF Health Byrd Alzheimer's Institute, University of South Florida, Tampa, Florida, USA
170 Fred Hutchinson Cancer Research Center, Seattle, Washington, USA.
171 Department of Molecular Neuroscience, UCL, Institute of Neurology, London, UK.
172 Mental Health & Behavioral Science Service, Bruce W. Carter VA Medical Center, Miami, FL
173 Department of Medical Sciences, Institute of Biomedicine iBiMED, University of Aveiro, 3810-193 Aveiro, Portugal
174 Department of Neuroscience, Mayo Clinic, Jacksonville, FL, USA.
175 Department of Neurology, University of Southern California, Los Angeles, California, USA.
176 Department of Neurodegenerative Disease, UCL Institute of Neurology, London, UK
177 Department of Neurology, Catholic University of Rome, Rome, Italy.
178 Department of Psychiatry and Behavioral Sciences, Miller School of Medicine, University of Miami, Miami, Florida, USA.
179 Department of Neurology, University of California, Davis, Sacramento, California, USA.
180 Institute for Memory Impairments and Neurological Disorders, University of California, Irvine, Irvine, California, USA.
181 Wien Center for Alzheimer's Disease and Memory Disorders, Mount Sinai Medical Center, Miami Beach, Florida, USA.
182 Rush Institute for Healthy Aging, Department of Internal Medicine, Rush University Medical Center, Chicago, Illinois, USA.
183 Department of Old Age Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.
184 Department of Primary Medical Care, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany.
185 Department of Pathology, University of Alabama at Birmingham, Birmingham, Alabama, USA.

186 Sanders-Brown Center on Aging, Department of Biostatistics, University of Kentucky, Lexington, Kentucky, USA.
187 Department of Psychiatry, New York University, New York, New York, USA.
188 Department of Psychiatry and Psychotherapy, University of Cologne, Cologne, Germany.
189 Department of Psychiatry and Psychotherapy, Charité University Medicine, Berlin, Germany.
190 Department of Neurosciences, University of California, San Diego, La Jolla, California, USA.
191 Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia, USA.
192 Emory Alzheimer's Disease Center, Emory University, Atlanta, Georgia, USA.
193 Department of Psychiatry, University of Freiburg, Freiburg, Germany (M.H.).
194 Neurogenetics Program, University of California, Los Angeles, Los Angeles, California, USA.
195 Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, USA.
196 Hope Center Program on Protein Aggregation and Neurodegeneration, Washington University School of Medicine, St. Louis, Missouri, USA.
197 Division of Genetics, Department of Medicine and Partners Center for Personalized Genetic Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA.
198 Department of Neurology, Washington University, St. Louis, Missouri, USA.
199 Department of Genetics, Washington University, St. Louis, Missouri, USA.
200 Department of Neurology, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts, USA.
201 Department of Pathology (Neuropathology), University of Pittsburgh, Pittsburgh, Pennsylvania, USA.
202 Department of Neurology, University of Alabama at Birmingham, Birmingham, Alabama, USA.
203 Cologne Center for Genomics, University of Cologne, Cologne, Germany.
204 UKDRI Cardiff, Cardiff University, Cardiff, UK.
205 Department of Pathology, Duke University, Durham, North Carolina, USA.
206 Department of Genome Sciences, University of Washington, Seattle, Washington, USA.
207 Department of Medicine (Medical Genetics), University of Washington, Seattle, Washington, USA.
208 Sanders-Brown Center on Aging, College of Public Health, Department of Epidemiology, University of Kentucky, Lexington, Kentucky, USA.
209 Department of Pathology and Laboratory Medicine, University of California, Davis, Sacramento, California, USA.
210 Unidad Clínica de Enfermedades Infecciosas y Microbiología. Hospital Universitario de Valme, Sevilla.
211 Department of Medicine (Biomedical Genetics), Boston University School of Medicine, Boston, MA, USA.
212 Functional Genomics Center Zurich, ETH/University of Zurich, Switzerland.
213 Department of Neurology, Oregon Health &Science University, Portland, Oregon, USA.
214 Department of Neurology, Portland Veterans Affairs Medical Center, Portland, Oregon, USA.
215 Department of Pathology and Laboratory Medicine, University of California, Irvine, Irvine, California, USA.
216 Department of Pathology, Boston University School of Medicine, Boston University, Boston, Massachusetts, USA.
217 INSERM U558, University of Toulouse, Toulouse, France.
218 Department of Neuropsychology, University of California San Francisco, San Francisco, California, USA.
219 Institute for Ageing and Health, Newcastle University, Biomedical Research Building, Campus for Ageing and Vitality, Newcastle upon Tyne.
220 Department of Neurobiology and Behavior, University of California, Irvine, Irvine, California, USA.
221 Department of Neurology, Emory University, Atlanta, Georgia, USA.
222 Institute for Stroke and Dementia Research, Klinikum der Universität München, Munich, Germany.
223 German Center for Neurodegenerative Diseases (DZNE), Munich, Germany.
224 Cleveland Clinic Lou Ruvo Center for Brain Health, Cleveland Clinic, Cleveland, Ohio, USA.
225 Division of Neuroscience and Experimental Psychology, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester M13 9PT, UK.
226 Department of Pathology, University of Michigan, Ann Arbor, Michigan, USA.

227 Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.
228 Helmholtz Center Munich, Institute of Epidemiology, Neuherberg.
229 Ludwig-Maximilians University Chair of Epidemiology, Munich, Germany.
230 Joint Biobank Munich and KORA Biobank.
231 Department of Psychiatry, Johns Hopkins University, Baltimore, Maryland, USA.
232 Human Genetics, Schools of Life Sciences and Medicine, University of Nottingham, Nottingham, UK.
233 Department of Medicine-Pulmonary, New York University, New York, New York, USA.
234 Department of Neurology, University of Miami, Miami, Florida, USA.
235 Department of Pathology, University of California, San Diego, La Jolla, California, USA.
236 Institute of Neurology, Catholic University of Sacred Hearth, Rome, Italy.
237 School of Nursing Northwest Research Group on Aging, University of Washington, Seattle, Washington, USA.
238 Institute of Primary Care and Public Health, Cardiff University, Neuadd Meirionnydd, University Hospital of Wales, Heath Park, Cardiff UK.
239 Department of Neurology, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA.
240 Institute of Public Health, University of Cambridge, Cambridge, UK.
241 Weill Institute for Neurosciences, Memory and Aging Center, University of California, San Francisco, San Francisco, CA, USA
242 Institute of Social Medicine, Occupational Health and Public Health, University of Leipzig, Leipzig, Germany.
243 Department of Pathology, University of Southern California, Los Angeles, California, USA.
244 King’s College London, Institute of Psychiatry, Psychology and Neuroscience, UK.
245 UK Dementia Research Institute at King's College London, UK.
246 LVR-Hospital Essen, Department of Psychiatry and Psychotherapy, University Duisburg-Essen, Germany.
247 Molecular Genetics Laboratory-Hospital, University of Central Asturias, Oviedo, Spain.
248 Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA.
249 Molecular Genetics Section, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA.
250 Department of Neurology, University of California, Irvine, Irvine, California, USA.
251 NEUROFARBA (Department of Neuroscience, Psychology, Drug Research and Child Health), University of Florence, Florence, Italy.
252 Centro di Ricerca, Trasferimento e Alta Formazione DENOTHE, University of Florence, Florence, Italy.
253 Department of Neurology, University of Colorado School of Medicine, Aurora, Colorado, USA.
254 Department of Neuroscience, Psychology, Drug Research and Child Health, University of Florence, Italy Viale Pieraccini 6, 50139 Florence, Italy.
255 IRCCS 'Don Carlo Gnocchi', Florence, Italy.
256 Alzheimer's Disease Center, New York University, New York, New York, USA.
257 Oxford Healthy Aging Project (OHAP), Clinical Trial Service Unit, University of Oxford, Oxford, UK.
258 Department of Epidemiology, Columbia University, New York, New York, USA.
259 Oxford Project to Investigate Memory and Ageing (OPTIMA), University of Oxford, Level 4, John Radcliffe Hospital, Oxford OX3 9DU, UK.
260 Department of Neurology, University of California, Los Angeles, Los Angeles, California, USA.
261 Department of Neurology, University of Texas Southwestern Medical Center, Dallas, Texas, USA.
262 Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK.
263 Department of Pathology (Neuropathology), Rush University Medical Center, Chicago, Illinois, USA.
264 Department of Psychiatry, University of Southern California, Los Angeles, California, USA.
265 Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA.
266 Department of Pathology, Johns Hopkins University, Baltimore, Maryland, USA.

267 Sanders-Brown Center on Aging, Department of Neuroscience, University of Kentucky, Lexington, Kentucky, USA.
268 Department of Pathology and Laboratory Medicine, University of California, Los Angeles, Los Angeles, California, USA.
269 Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Pathology, Columbia University, New York, New York, USA.
270 Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA.
271 Department of Psychiatry and Behavioral Sciences, Duke University, Durham, North Carolina, USA.
272 Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA.
273 Department of Pathology, Oregon Health &Science University, Portland, Oregon, USA.
274 Evelyn F. McKnight Brain Institute, Department of Neurology, Miller School of Medicine, University of Miami, Miami, Florida, USA.
275 Institute of Molecular Biology and Biochemistry, Centre for Molecular Medicine, Medical University Graz, Graz, Austria.
276 Gerontology and Geriatrics Research Laboratory, IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy.
277 IdiPAZ, Instituto de Investigación Sanitaria la Paz, Spain.
278 Centro de Biologia Molecular Severo Ochoa (CSIC-UAM), Madrid, Spain.
279 German Center for Neurodegenerative Diseases (DZNE), Berlin, Germany.
280 Instituto di Ricovero e Cura a Carattere Scientifico (IRCCS) Associazione Oasi Maria Santissima Srl, Troina, Italy.
281 Center for Translational and Computational Neuroimmunology, Department of Neurology, Columbia University Medical Center, New York, New York, USA.
282 Memory Unit, Neurology Department and Sant Pau Biomedical Research Institute, Hospital de la Santa Creu i Sant Pau, Autonomous University Barcelona, Barcelona, Spain.
283 Somerset Partnership NHS Trust, Somerset, UK.
284 The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK.
285 Theme Aging, Genetics Unit, Karolinska University Hospital Huddinge, S-14186 Stockholm; 14, Karolinska Institutet, Dept Neurobiology, Care Sciences and Society, Alzheimer Research Center, Novum floor 5, S-141 57 Huddinge, Sweden.
286 Karolinska Institutet, Dept Neurobiology, Care Sciences and Society, Alzheimer Research Center, Novum floor 5, S-141 57 Huddinge, Sweden.
287 Ronald M. Loeb Center for Alzheimer's Disease, Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, New York, USA.
288 ~~University of Bristol Institute of Clinical Neurosciences, School of Clinical Sciences, Frenchay Hospital, Bristol, UK.~~
289 Department of Neurology, Mayo Clinic, Jacksonville, Florida, USA.
290 Memory Research and Resources Center, CMRR de Bordeaux, Bordeaux, France.
291 Utah State University, Logan, Utah, USA.
292 Department of Neurology, Erasmus MC University Medical Center, Rotterdam, the Netherlands.
293 Departments of Radiology, Erasmus MC University Medical Center, Rotterdam, The Netherlands.
294 Department of Neurology, Medical University Graz, Graz, Austria.
295 Center for Applied Genomics, Children's Hospital of Philadelphia, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA.
296 Division of Human Genetics, Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA.
297 Department of Ophthalmology, Boston University School of Medicine, Boston University, Boston, Massachusetts, USA.
298 Department of Epidemiology, Boston University School of Public Health, Boston, Massachusetts, USA.
299 Department of Health Services, University of Washington, Seattle, Washington, USA.
300 Group Health Research Institute, Group Health, Seattle, Washington, USA.
301 Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases, San Antonio, TX, USA.

+equal contribution first author
~equal contribution senior author
*corresponding author

Materials and Correspondence:

Brian W. Kunkle, PhD, MPH
Hussman Institute for Human Genomics
Miller School of Medicine
University of Miami
1501 NW 10th Ave
Miami, FL 33136
bkunkle@miami.edu

Jean-Charles Lambert, PhD
INSERM, U1167
Laboratoire d'Excellence Distalz
Institut Pasteur de Lille
University Lille
F59000 Lille, France
jean-charles.lambert@pasteur-lille.fr

Margaret Pericak-Vance, PhD
Hussman Institute for Human Genomics
Miller School of Medicine
University of Miami
1501 NW 10th Ave
Miami, FL 33136
mpericak@miami.edu

Introduction

Risk for Late-onset Alzheimer's disease (LOAD), the most prevalent dementia in the elderly[1], is partially driven by genetics[2]. To identify LOAD risk loci, we performed the largest genome-wide association meta-analysis of clinically diagnosed LOAD to date (94,437 individuals), analyzing both common and rare variants. We confirm 20 previous LOAD risk loci and identify five new genome-wide loci (*IQCK*, *ACE*, *ADAM10*, *ADAMTS1* and *WWOX*). Fine-mapping of the human leukocyte antigen (HLA) region confirms the neurological and immune-mediated disease haplotype HLA-DR15 as a risk factor for LOAD. Pathway analysis implicates the immune system and lipid metabolism, and for the first time tau binding proteins and APP metabolism, showing that genetic variants affecting APP and Aβ processing are not only associated with early-onset autosomal dominant AD but also with LOAD. Analysis of AD risk genes and pathways show enrichment for rare variants ($P=1.32\times10^{-7}$) indicating that additional rare variants remain to be identified. Finally, we also identify important genetic correlations between LOAD and other traits including family history of dementia and education.

Main Text

Our previous work identified 19 genome-wide significant common variant signals in addition to *APOE*[3], that influence risk for LOAD (onset age > 65 years). These signals, combined with 'subthreshold' common variant associations, account for ~31% of the genetic variance of LOAD[2], leaving the majority of genetic risk uncharacterized[4]. To search for additional signals, we conducted a genome-wide association (GWAS) meta-analysis of non-Hispanic Whites (NHW) using a larger sample (17 new, 46 total datasets) from our group, the International Genomics of Alzheimer's Project (IGAP) (composed of four AD consortia: ADGC, CHARGE, EADI, and GERAD). This sample increases our previous discovery sample (Stage 1) by 29% for cases and 13% for controls (N=21,982 cases; 41,944 controls) (**Supplementary Tables 1**

and **2**, and **Supplementary Note**). To sample both common and rare variants (minor allele frequency MAF ≥ 0.01, and MAF < 0.01, respectively), we imputed the discovery datasets using a 1000 Genomes reference panel consisting of 36,648,992 single-nucleotide variants, 1,380,736 insertions/deletions, and 13,805 structural variants. After quality control, 9,456,058 common variants and 2,024,574 rare variants were selected for analysis (a 63% increase from our previous common variant analysis in 2013). Genotype dosages were analyzed within each dataset, and then combined with meta-analysis (**Supplementary Figure 1, Supplementary Tables 1-3**).

The Stage 1 discovery meta-analysis produced 12 loci with genome-wide significance ($P \leq 5 \times 10^{-8}$) (**Table 1**), all of which are previously described[3,5–12]. Genomic inflation factors were slightly inflated (lambda median=1.05; lambda regression=1.09; See **Supplementary Figure 2** for QQ-plot), however, univariate linkage disequilibrium score (LDSC) regression[13,14] estimates indicated that the majority of this inflation was due to a polygenic signal, with the intercept being close to 1 (1.026, s.e. = 0.006). The observed heritability (h^2) of LOAD was estimated at 0.071 (0.011) using LDSC. Stage 1 meta-analysis was first followed by Stage 2 using the I-select chip we previously developed in Lambert et al.[3] (including 11,632 variants, N=18,845; **Supplementary Table 4**) and finally Stage 3A (N=11,666) or Stage 3B (N=30,511) (for variants in regions not well captured in the I-select chip) (See **Supplementary Figure 1** for workflow). The final sample was 35,274 clinical and autopsy-documented AD cases and 59,163 controls.

Meta-analysis of Stages 1 and 2 produced 21 genome-wide significant associations ($P \leq 5 \times 10^{-8}$) (**Table 1** and **Figure 1**). Of these, 18 were previously reported as genome-wide significant in Lambert et al.[3]. Three other signals were not initially described in the initial IGAP GWAS: the rare R47H *TREM2* coding variant previously reported by others[8,9,15]; *ECDH3* (rs7920721) which was recently identified as a potential genome-wide significant AD risk locus in several studies[23-25] and *ACE* (rs138190086) (**Supplementary Figures 3-4**). In addition, seven signals showed suggestive association with a $P < 5 \times 10^{-7}$ (respectively rs593742,

rs830500, rs7295246, rs7185636, rs2632516, rs4735340, and rs10467994 for their closest gene *ADAM10, ADAMTS1, ADAMTS20, IQCK*, *miR142/TSPOAP1-AS1*, *NDUFAF6* and *SPPL2A*) (**Supplementary Figures 5-11**). Stage 3A and meta-analysis of all three stages for these 9 variants (excluding the *TREM2* signal; see **Supplementary Table 5** for variant list) identified five genome-wide significant loci. In addition to *ECDH3*, this included four new genome-wide AD risk signals not previously described in other clinical AD GWAS at *IQCK, ADAMTS1, ACE* and *ADAM10* (**Table 2**)*. ACE* and *ADAM10* were previously reported as AD candidate genes[16–20] that were not replicated in some subsequent studies[19,21–24]. Two of the four other signals were close to genome-wide significance: *miR142/TSPOAP1-AS1* ($P = 5.3 \times 10^{-8}$) and *NDUFAF6* ($P = 9.2 \times 10^{-8}$) (**Table 2**). We also extended the analyses of the two loci (*NME8* and *MEF2C*) in Stage 3 that were previously genome-wide significant in our 2013 meta-analysis. These loci were not genome-wide significant in our current study and will deserve further investigations (*NME8*: $P = 2.7 \times 10^{-7}$; *MEF2C*: $P = 9.1 \times 10^{-8}$; **Supplementary Figures 12-13**). Of note, GCTA-COJO[25] conditional analysis of the genome-wide loci indicates that *TREM2* and three other loci (*BIN1*, *ABCA7*, and *PTK2B/CLU*) have multiple independent LOAD association signals (**Supplementary Table 6**), suggesting that the genetic variance associated with some GWAS loci is probably under-estimated.

We also selected 33 variants from Stage 1 (28 common variants and 5 rare variants in loci not well captured in the I-select chip; see methods for full selection criteria) for genotyping in Stage 3B (including populations of Stage 2 and Stage 3A). We nominally replicated a rare variant (rs71618613) within an intergenic region near *SUCLG2P4* (MAF = 0.01; $P = 6.8 \times 10^{-3}$; combined-$P = 3.3 \times 10^{-7}$) and replicated a low-frequency variant in the *TREM2* region (rs114812713, MAF=0.03, $P = 7.2 \times 10^{-3}$; combined-$P = 2.1 \times 10^{-13}$) in the gene *OARD1* that may represent an independent signal according to our conditional analysis (**Table 2, Supplementary Figures 14-15, Supplementary Tables 6 and 7**). In addition, rs62039712 in the *WWOX* locus reached genome-wide significance ($P = 3.7 \times 10^{-8}$) and rs35868327 in the *FST* locus reached

suggestive significance ($P = 2.6 \times 10^{-7}$) (**Table 2 and Supplementary Figures 16-17**). *WWOX* may play a role in AD through its interaction with Tau[26,27], and it's worth noting the sentinel variant (defined as variants with the lowest p-values) is just 2.4 megabases from *PLCG2*, which contains a rare variant we recently associated with AD[15]. Since both rs62039712 and rs35868327 were only analyzed in a restricted number of samples, these loci deserve further attention.

To evaluate the biological significance and attempt to identify the underlying risk genes for the newly identified genome-wide signals (*IQCK*, *ACE*, *ADAM10*, *ADAMTS1* and *WWOX*) and those found previously, we pursued five strategies: 1) annotation and gene-based testing for deleterious coding, loss-of-function (LOF) and splicing variants, 2) expression-quantitative trait loci (eQTL) analyses, 3) evaluation of transcriptomic expression in LOAD clinical traits (correlation with BRAAK stage[28] and differential expression in AD versus control brains[29]), 4) evaluation of transcriptomic expression in AD-relevant tissues[30–32], and 5) gene cluster/pathway analyses. For the 24 signals reported here, other evidence indicates that *APOE*[33,34], *ABCA7*[35–38], *BIN1*[39], *TREM2*[8,9], *SORL1*[40,41], *ADAM10*[42], *SPI1*[43], *and CR1*[44] are the true AD risk gene, though there is a possibility that multiple risk genes exist in these regions[45]. Because many GWAS loci are intergenic, and the closest gene to the sentinel variant may not be the actual risk gene, in these analyses, we considered all protein coding genes within ±500kb of the sentinel variant linkage disequilibrium (LD) regions ($r^2 \geq 0.5$) for each locus as a candidate AD gene (N = 400 genes) (**Supplementary Table 8**).

We first annotated all sentinel variants for each locus and variants in LD ($r^2 > 0.7$) with these variants in a search for deleterious coding, loss-of-function (LOF) or splicing variants. In line with findings that most causal variants for complex disease are non-coding[46], only 2% of 1,073 variants across the 24 loci (excluding *APOE*) were exonic variants, with a majority (58%) being intronic (**Supplementary Figure 18** and **Supplementary Table 9**). Potentially deleterious variants include the rare R47H missense variant in *TREM2*, common missense variants in *CR1*,

SPI1, *MS4A2*, and *IQCK*, and a relatively common (MAF = 0.16) splicing variant in *IQCK*. Using results of a large whole-exome sequencing study conducted in the ADGC and CHARGE sample[47] (N = 5,740 LOAD cases and 5,096 cognitively normal controls), we also identified 10 genes located in our genome-wide loci as having rare deleterious coding, splicing or LOF burden associations with LOAD (FDR $P < 0.01$), including previously implicated rare-variant signals in *ABCA7*, *TREM2*, and *SORL1*[15,47–53], and additional associations with *TREML4* in the *TREM2* locus, *TAP2* and *PSMB8* in the *HLA-DRB1* locus, *PIP* in the *EPHA1* locus, *STYX* in the *FERMT2* locus, *RIN3* in the *SLC24A4* locus, and *KCNH6* in the *ACE* locus (**Supplementary Table 10**).

For eQTL analyses, we searched existing eQTL databases and studies for cis-acting eQTLs in a prioritized set of variants (N = 1,873) with suggestive significance or in LD with the sentinel variant in each locus. 71-99% of these variants have regulatory potential when considering all tissues according to RegulomeDB[54] and HaploReg[55], but restricting to AD-relevant tissues (via Ensembl Regulatory Build[56] and GWAS4D[57]) appears to aid in regulatory variant prioritization, with probabilities for functional variants increasing substantially when using GWAS4D cell-dependent analyses with brain or monocytes for instance (these and other annotations are provided in **Supplementary Table 11**). Focusing specifically on eQTLs, we found overlapping cis-acting eQTLs for 153 of the 400 protein coding genes, with 136 eQTL-controlled genes in AD relevant tissues (i.e. brain and blood/immune cell types; see methods for details) **(Supplementary Tables 12 and 13**). For our newly identified loci, there were significant eQTLs in AD relevant tissue for: *ADAM10* in prefrontal cortex and blood, *FAM63B* in blood, and *SLTM* in putamen in the *ADAM10* locus; *ADAMTS1* in blood in the ADAMTS1 locus; and *ACSM1* and *ANKS4B* in monocytes, *C16orf62* in blood, *GDE1* in cerebellum, and *GPRC5B*, *IQCK*, and *KNOP1* in several brain and blood tissue types in the *IQCK* locus. There were no eQTLs in AD-relevant tissues in the *WWOX* or *ACE* locus, though several eQTLs for *PSMC5* in coronary artery tissue were found for the *ACE* locus. eQTL's for genes in previously identified

loci include *BIN1* in monocytes and cerebellum, *INPP5D* in prefrontal cortex and blood, *CD2AP* in cerebellum and prefrontal cortex, and *SLC24A4* in monocytes. Co-localization analysis confirmed evidence of a shared causal variant affecting expression and disease risk in 66 genes over 20 loci, including 31 genes over 13 loci in LOAD relevant tissue (see **Supplementary Table 14 and 15** for a complete lists). Genes implicated include: *CR1* and *ABCA7* in brain (in the *CR1* and *ABCA7* loci respectively); *BIN1* (in the *BIN1* locus), *SPI1* and *MYBPC3* (both in the *SPI1* locus) in blood; *MS4A2*, *MS4A6A*, and *MS4A4A* (all at the *MS4A2* locus) in blood; and *KNOP1* (in the *IQCK* locus) and *HLA-DRB1* (in the *HLA-DRB1* locus) in both blood and brain. (**Supplementary Table 12**).

To study the differential expression of genes in brains of AD patients versus controls, we used thirteen expression studies[29]. 58% of the 400 protein coding genes within the genome-wide loci had evidence of differential expression in at least one study (**Supplementary Table 16**). Additional comparisons to AD related gene expression sets revealed 62 genes were correlated with pathogenic stage (BRAAK) in at least one brain tissue[28] (44 genes in the prefrontal cortex, the most relevant LOAD tissue; 36 in cerebellum, and 1 in visual cortex). Finally, 38 genes were present in a set of 1,054 genes preferentially expressed in aged microglial cells, a gene set shown to be enriched for AD genes ($P = 4.1 \times 10^{-5}$)[32]. We also annotated our list of genes with Brain RNA-seq data which showed that 80% were expressed in at least one type of brain cell, and the genes were most highly expressed in fetal astrocytes (26%), followed by microglia/macrophage (15.8%), neurons (14.8%), astrocytes (11.5%) and oligodendrocytes (6.5%). When not considering fetal astrocytes, mature astrocytes (21%) and microglial cells (20.3%), the resident macrophage cell of the brain thought to play a key role in the pathologic immune response in LOAD[9,15,58], become the highest expressed cell type (20.3%) in the genome-wide set of genes, with 5.3% of the 400 genes showing high microglial expression (**Supplementary Table 17**; see **Supplementary Table 18** for highly expressed gene list by cell type).

We conducted pathway analyses (MAGMA[59]) using five gene set resources. Analysis were conducted separately for common (MAF $\geq$ 0.01) and rare variants (MAF < 0.01). For common variants, we detected four function clusters including: 1) APP metabolism/Aβ-formation (regulation of beta-amyloid formation: $P = 4.56 \times 10^{-7}$ and regulation of amyloid precursor protein catabolic process: $P = 3.54 \times 10^{-6}$), 2) tau protein binding ($P = 3.19 \times 10^{-5}$), 3) lipid metabolism (four pathways including protein-lipid complex assembly: $P = 1.45 \times 10^{-7}$), and 4) immune response ($P = 6.32 \times 10^{-5}$) (**Table 3** and **Supplementary Table 19**). Enrichment of the four pathways remains after removal of genes in the *APOE* region. When *APOE*-region genes and genes in the vicinity of genome-wide significant genes are removed, tau shows moderate association ($P = 0.027$) and lipid metabolism and immune related pathways show strong associations ($P < 0.001$) (**Supplementary Table 20**). Genes driving these enrichments (i.e. having a gene-wide $P < 0.05$) include *SCNA*, a Parkinson's risk gene that encodes alpha-synuclein, the main component of Lewy bodies, and may play a role in tauopathies[60,61], for the tau pathway; apolipoprotein genes (*APOM*, *APOA5*) and *ABCA1*, a major regulator of cellular cholesterol, for the lipid metabolism pathways; and 52 immune pathway genes (**Supplementary Table 21**). While no pathways were significantly enriched for rare variants, lipid and Aβ-pathways did have nominal significance in rare-variant-only analyses. Importantly, we also observe a highly significant correlation between common and rare pathway gene results ($P = 1.32 \times 10^{-7}$), suggesting that risk AD genes and pathways are enriched for rare variants. In fact, 50 different genes within tau, lipid, immunity and Aβ pathways show nominal rare-variant driven associations ($P < 0.05$) with LOAD.

To further explore the APP/Aβ-pathway enrichment we analyzed a comprehensive set of 335 APP metabolism genes[62] curated from the literature. We observed significant enrichment of this gene-set in common variants ($P = 2.27 \times 10^{-4}$; $P = 3.19 \times 10^{-4}$ excluding *APOE*), with both *ADAM10* and *ACE* nominally significant drivers of this result (**Table 4** and **Supplementary Table 22 and 23**). Several 'sub-pathways' were also significantly enriched in the common-

variants including ‘clearance and degradation of Aβ’ and ‘aggregation of Aβ’, along with its subcategory ‘microglia’, the latter supporting microglial cells suspected role in response to Aβ in LOAD[63]. Nominal enrichment for risk from rare variants was found for the pathway ‘aggregation of Aβ: chaperone’ and 23 of the 335 genes.

To identify candidate genes for our novel loci, we combined results from our eQTL, clinical and AD-relevant tissue expression, and gene function/pathway analyses in a priority ranking method similar to Fritsche et al.[64] (**Table 5 and Supplementary Table 24**). For our *ADAM10* signal, of the 11 genes within this locus, *ADAM10* was the top ranked gene. *ADAM10*, the most important α-secretase in the brain, is a component of the non-amyloidogenic pathway of APP metabolism[65], and sheds *TREM2*[66]*,* an innate immunity receptor expressed selectively in microglia. Over-expression of *ADAM10* in mouse models can halt Aβ production and subsequent aggregation[67]. Also two rare *ADAM10* mutations segregating with disease in LOAD families increased Aβ plaque load in “Alzheimer-like” mice, with diminished α-secretase activity from the mutations likely the causal mechanism[17,42]. For the *IQCK* signal, which is also an obesity locus[68,69], *IQCK*, a relatively uncharacterized gene, was top ranked, though four of the other 11 genes in the locus have a priority rank ≥ 4 including *KNOP1* and *GPRC5B,* the latter being a regulator of neurogenesis[70,71] and inflammatory signalling in obesity[72]. Of the 22 genes in the *ACE* locus, *PSMC5*, a key regulator of major histocompatibility complex (MHC)[73,74], has a top score of 4, while *DDX42*, *MAP3K3*, an important regulator of macrophages and innate immunity[75,76], and *CD79B,* a B lymphocyte antigen receptor sub-unit each have a score of 3. Candidate gene studies previously associate *ACE* variants with AD risk[18,20,77], including a strong association in the Wadi Ara, an Israeli Arab community with high risk of AD[19]. However, these studies yielded inconsistent results[21], and our work is the first to report a clear genome-wide association in NHW at this locus. While our analyses did not prioritize *ACE,* it should not be rejected as a candidate gene, as its expression in AD brain tissue is associated with Aβ load and AD severity[78]. Furthermore, CSF levels of the angiotensin-converting enzyme (ACE) are

associated with Aβ levels[79] and LOAD risk[80], and studies show ACE can inhibit Aβ toxicity and aggregation[81]. Finally, angiotensin II, a product of ACE function mediates a number of neuropathological processes in AD[82] and is now a target for intervention in phase II clinical trials of AD[83]. Another novel genome-wide locus reported here *ADAMTS1,* is within 665 kb of *APP* on chromosome 21. Of three genes at this locus (*ADAMTS1*, *ADAMTS5*, *CYYR1*), our analyses nominates *ADAMTS1,* as the likely risk gene, though we cannot rule out that this signal is a regulatory element for *APP*. *ADAMTS1* is elevated in Down Syndrome with neurodegeneration and AD[84] and is a potential neuroprotective gene[85,86,87], or a neuroinflammatory gene important to microglial response[88]. Finally, *WWOX* and *MAF*, which surround an intergenic signal in an obesity associated locus[89], were both prioritized, with *MAF,* another important regulator of macrophages[90,91], being highly expressed in microglia in the Brain-RNA-seq database, and *WWOX*, an HDL-cholesterol and triglyceride associated gene[92,93], being expressed in several brain cell types, most highly in astrocytes and neurons. *WWOX* has been implicated in several neurological phenotypes[94], binds Tau and may play a critical role in regulating Tau hyper-phosphorylation, neurofibrillary formation, and amyloid β aggregation[26,27]. Intriguingly, treatment of mice with its binding partner restores memory deficits[95], hinting at its potential in neurotherapy.

For previously reported loci, named for the closest gene, applying the same approach for prioritization highlights several genes as described in **Table 5**, some of which are involved in APP metabolism (*FERMT2, PICALM*) or Tau toxicity (*BIN1*, *CD2AP*, *FERMT2*, *CASS4*, *PTK2B*)[96–99]. Pathway, tissue and disease traits enrichment analysis supports the utility of our prioritization method, as the 53 prioritized genes with a score ≥ 5 are: 1) enriched in substantially more AD relevant pathways, processes, and dementia-related traits, 2) enriched in candidate AD cell types such as monocytes (adjusted-P = 9.0 x 10^{-6}) and macrophages (adjusted-P = 5.6 x 10^{-3}), and 3) increased in strength of associations for dementia-related traits and AD relevant pathways (**Supplementary Table 25 and 26;** see **Supplementary Figure 19**

for the interaction network of these prioritized genes). To further investigate the cell types and tissues the prioritized genes are expressed in, we performed differentially expressed gene (DEG) set enrichment analysis of the prioritized genes using GTEx[100] tissues, and identified significant differential expression in several potentially relevant AD tissues including: immune-related tissues (upregulation in blood and spleen), obesity-related tissue (upregulation in adipose), heart tissues (upregulation in left ventricle and atrial appendage) and brain tissues (dowregulation in cortex, cerebellum, hippocampus, basal ganglia, and amygdala). Furthermore, the 53 genes are overexpressed in 'adolescence' and 'young adult' brain tissues in BrainSpan[101], a transcriptomics atlas of the developing human brain, which is consistent with accumulating evidence suggesting AD may start decades before the onset of disease[102,103] (**Supplementary Figure 20;** see **Supplementary Figure 21** for a tissue expression heat map for the 53 genes).

The above approach prioritized *HLA-DRB1* as the top candidate gene in the MHC locus, known for its complex genetic organization and highly polymorphic nature (see **Supplementary Figure 22** for Stage 1 results plot of region). Previous analyses in the ADGC (5,728 AD cases and 5,653 controls) has linked both HLA class I and class II haplotypes with AD risk[104]. In order to further investigate this locus in a much larger sample, we used a robust imputation method and fine-mapping association analysis of alleles and haplotypes of HLA class I (*HLA-A*, *HLA-B*, *HLA-C*) and class II (*HLA-DRB1*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DPA1*, *HLA-DPB1*) genes in 14,776 cases and 23,047 controls from our datasets (**Supplementary Table 27**) (see Methods). We found risk effects of *HLA-DQA1*01:02* (false discovery rate (FDR) $P = 0.014$), *HLA-DRB1*15:01* (FDR $P = 0.083$) and *HLA-DQB1*06:02* (FDR $P = 0.010$) (**Supplementary Table 28**). After conditioning on the sentinel variant in this region from the meta-analysis (rs78738018), association signals were lost for the three alleles suggesting that the signal observed at the variant level is due to the association of these three alleles. These alleles form the *HLA-DQA1*01:02~HLA-DQB1*06:02~HLA-DRB1*15:01* (*DR15*) haplotype, which is also

associated with AD in our sample (FDR *P* = 0.013) (**Supplementary Table 29**). When considering only 2-loci haplotypes, *HLA-DQB1*06:02~HLA-DRB1*15:01* (FDR *P* = 0.013), *HLA-DQA1*01:02~HLA-DRB1*15:01* (FDR *P* = 0.013), *HLA-DQA1*01:02~HLA-DQB1*06:02* (FDR *P* = 0.013) also show association with AD. Taken together, these results suggest a central role of the HLA-*DQA1*01:02~HLA-DQB1*06:02~HLA-DRB1*15:01* haplotype in AD risk. This haplotype was associated with risk of AD originally in a small study in the Tunisian population[105], and more recently in a large ADGC analysis[104]. Intriguingly, this haplotype and its component alleles also associate with protection against diabetes[106], a high risk for multiple sclerosis[107,108], and risk or protective effects with many other immune-mediated diseases (**Supplementary Table 30**). Moreover, the associated diseases at these loci include a large number of traits queried from an HLA-specific Phewas[109], including neurological diseases (i.e. Parkinson's disease[110,111]) and diseases with risk factors for AD (i.e. hyperthyroidism[112]), pointing to potential shared and/or interacting mechanisms and co-morbidities, a common paradigm in the MHC locus[113]. Two additional alleles, *HLA-DQA1*03:01* and *HLA-DQB1*03:02*, belonging to another haplotype, show protective effect on AD, but their signal was lost after conditioning on *HLA-DQA1*01:02* and the *HLA-DQA1*03:01~HLA-DQB1*03:02* haplotype is not associated with AD (FDR *P* = 0.651).

As described above, several of our genome-wide loci have potentially interesting co-morbid or pleiotropic associations with traits that may be relevant to pathology of AD. To investigate the extent of LOAD's shared genetic architecture with other traits we performed LD-score regression to estimate the genetic correlation between LOAD and 792 human diseases, traits, and behaviors[13,114] (**Supplementary Table 31**). The common variant genetic architecture of LOAD was positively correlated with maternal family history of Alzheimer's disease/dementia (rg = 0.81; FDR $P = 2.79 \times 10^{-7}$), similar to a recent GWAS using family history of AD as a proxy[115] which found maternal genetic correlation with AD to be higher than paternal AD (rg = 0.91 and 0.66 respectively). There is substantial overlap between these estimates as the

Marioni et al. analyses include the 2013 IGAP summary statistics and employed the same UK Biobank variable that we used for rg estimates with maternal history of dementia. While use of proxy AD cases introduces less sensitivity and specificity for true AD signals overall in comparison to clinically-diagnosed AD association analyses, the investigation did identify 17 of our 25 genome-wide loci including the *ACE* and *ADAM10* loci, suggesting that familial proxy AD studies can identify AD relevant loci. We also find significant negative correlation between AD and multiple measures of educational attainment (i.e. college completion, rg = -0.24; years of schooling, rg range = -0.19 to -0.24; cognitive scores, rg's = -0.24 and -0.25) (FDR $P < 0.05$), supporting the theory that a greater cognitive reserve could help protect against development of LOAD[116]. The extent to which socioeconomic (ses), environmental or cultural factors contribute to the correlation between educational attainment and risk for AD is unknown, but research has shown dementia risk to be associated with lower ses status, independent of education status[117,118]. Furthermore, we also found negative correlations at $P < 0.05$ with multiple measures of cardiovascular health (i.e. family history of high blood pressure, family history of heart disease, vascular/heart problems diagnosis) and diabetes (i.e. fasting proinsulin, basal metabolic rate, fasting insulin main effect), supporting previous research that suggested use of blood pressure and diabetic medications may reduce risk of AD[119]. In fact, use of blood pressure medication does show negative genetic correlation with AD in our study (rg = -0.12; $P = 0.035$), though this result does not survive FDR correction. These and other top results from this analysis (i.e. body mass index, height; see **Supplementary Table 31** for a full list of other nominally significant correlations) have been linked to AD previously[114,119–126], either through suggestive or significant genetic or epidemiological associations (see Kuzma et al. 2018[127] for a recent review), but the multiple measures here support and emphasize their genetic correlation with LOAD and highlight the possible genetic pleiotropy or co-morbidity of these traits with pathology of LOAD.

In conclusion, our work identifies five new genome-wide associations for LOAD and shows that GWAS data combined with high-quality imputation panels can reveal rare disease risk variants (i.e. *TREM2*). The enrichment of rare-variants in pathways associated with AD indicates that additional rare-variants remain to be identified, and larger samples and better imputation panels will facilitate identifying these rare variants. While these rare-variants may not contribute substantially to the predictive value of genetic findings, it will add to the understanding of disease mechanisms and potential drug targets. Discovery of the risk genes at genome-wide loci remains challenging, but we demonstrate that converging evidence from existing and new analyses can prioritize risk genes. We also show that APP metabolism is not only associated with early-onset but also late-onset AD, suggesting that therapies developed by studying early-onset families could also be applicable to the more common late-onset form of the disease. Pathway analysis showing tau is involved in late-onset AD supports recent evidence that tau may play an early pathological role in AD[128–130], and confirms that therapies targeting tangle formation/degradation could potentially affect late-onset AD. Finally, our fine-mapping analyses of HLA and genetic correlation results point to LOAD's shared genetic architecture with many immune-mediated and cognitive traits, and suggests that research and interventions that elucidate the mechanisms behind these relationships could also yield fruitful therapeutic strategies for LOAD.

Acknowledgements

ADGC. The National Institutes of Health, National Institute on Aging (NIH-NIA) supported this work through the following grants: ADGC, U01 AG032984, RC2 AG036528; Samples from the National Cell Repository for Alzheimer's Disease (NCRAD), which receives government support under a cooperative agreement grant (U24 AG21886) awarded by the National Institute on Aging (NIA), were used in this study. We thank contributors who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible; Data for this study were prepared, archived, and distributed by the National Institute on Aging Alzheimer's Disease Data Storage Site (NIAGADS) at the University of Pennsylvania (U24-AG041689-01); NACC, U01 AG016976; NIA LOAD (Columbia University), U24 AG026395, U24 AG026390, R01AG041797; Banner Sun Health Research Institute P30 AG019610; Boston University, P30 AG013846, U01 AG10483, R01 CA129769, R01 MH080295, R01 AG017173, R01 AG025259, R01 AG048927, R01AG33193, R01 AG009029; Columbia University, P50 AG008702, R37 AG015473, R01 AG037212, R01 AG028786; Duke University, P30 AG028377, AG05128; Emory University, AG025688; Group Health Research Institute, UO1 AG006781, UO1 HG004610, UO1 HG006375, U01 HG008657; Indiana University, P30 AG10133, R01 AG009956, RC2 AG036650; Johns Hopkins University, P50 AG005146, R01 AG020688; Massachusetts General Hospital, P50 AG005134; Mayo Clinic, P50 AG016574, R01 AG032990, KL2 RR024151; Mount Sinai School of Medicine, P50 AG005138, P01 AG002219; New York University, P30 AG08051, UL1 RR029893, 5R01AG012101, 5R01AG022374, 5R01AG013616, 1RC2AG036502, 1R01AG035137; North Carolina A&T University, P20 MD000546, R01 AG28786-01A1; Northwestern University, P30 AG013854; Oregon Health & Science University, P30 AG008017, R01 AG026916; Rush University, P30 AG010161, R01 AG019085, R01 AG15819, R01 AG17917, R01 AG030146, R01 AG01101, RC2 AG036650, R01 AG22018; TGen, R01 NS059873; University of Alabama at Birmingham, P50 AG016582; University of Arizona, R01 AG031581; University of California,

Davis, P30 AG010129; University of California, Irvine, P50 AG016573; University of California, Los Angeles, P50 AG016570; University of California, San Diego, P50 AG005131; University of California, San Francisco, P50 AG023501, P01 AG019724; University of Kentucky, P30 AG028383, AG05144; University of Michigan, P50 AG008671; University of Pennsylvania, P30 AG010124; University of Pittsburgh, P50 AG005133, AG030653, AG041718, AG07562, AG02365; University of Southern California, P50 AG005142; University of Texas Southwestern, P30 AG012300; University of Miami, R01 AG027944, AG010491, AG027944, AG021547, AG019757; University of Washington, P50 AG005136, R01 AG042437; University of Wisconsin, P50 AG033514; Vanderbilt University, R01 AG019085; and Washington University, P50 AG005681, P01 AG03991, P01 AG026276. The Kathleen Price Bryan Brain Bank at Duke University Medical Center is funded by NINDS grant # NS39764, NIMH MH60451 and by Glaxo Smith Kline. Support was also from the Alzheimer's Association (LAF, IIRG-08-89720; MP-V, IIRG-05-14147), the US Department of Veterans Affairs Administration, Office of Research and Development, Biomedical Laboratory Research Program, and BrightFocus Foundation (MP-V, A2111048). P.S.G.-H. is supported by Wellcome Trust, Howard Hughes Medical Institute, and the Canadian Institute of Health Research. Genotyping of the TGEN2 cohort was supported by Kronos Science. The TGen series was also funded by NIA grant AG041232 to AJM and MJH, The Banner Alzheimer's Foundation, The Johnnie B. Byrd Sr. Alzheimer's Institute, the Medical Research Council, and the state of Arizona and also includes samples from the following sites: Newcastle Brain Tissue Resource (funding via the Medical Research Council, local NHS trusts and Newcastle University), MRC London Brain Bank for Neurodegenerative Diseases (funding via the Medical Research Council),South West Dementia Brain Bank (funding via numerous sources including the Higher Education Funding Council for England (HEFCE), Alzheimer's Research Trust (ART), BRACE as well as North Bristol NHS Trust Research and Innovation Department and DeNDRoN), The Netherlands Brain Bank (funding via numerous sources including Stichting MS Research, Brain Net Europe, Hersenstichting Nederland Breinbrekend

Werk, International Parkinson Fonds, Internationale Stiching Alzheimer Onderzoek), Institut de Neuropatologia, Servei Anatomia Patologica, Universitat de Barcelona. ADNI data collection and sharing was funded by the National Institutes of Health Grant U01 AG024904 and Department of Defense award number W81XWH-12-2-0012. ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. We thank Drs. D. Stephen Snyder and Marilyn Miller from NIA who are *ex-officio* ADGC members.

EADI. This work has been developed and supported by the LABEX (laboratory of excellence program investment for the future) DISTALZ grant (Development of Innovative Strategies for a Transdisciplinary approach to ALZheimer's disease) including funding from MEL (Metropole européenne de Lille), ERDF (European Regional Development Fund) and Conseil Régional

Nord Pas de Calais. This work was supported by INSERM, the National Foundation for Alzheimer's disease and related disorders, the Institut Pasteur de Lille and the Centre National de Génotypage, the JPND PERADES, GENMED, and the FP7 AgedBrainSysBio. The Three-City Study was performed as part of collaboration between the Institut National de la Santé et de la Recherche Médicale (Inserm), the Victor Segalen Bordeaux II University and Sanofi-Synthélabo. The Fondation pour la Recherche Médicale funded the preparation and initiation of the study. The 3C Study was also funded by the Caisse Nationale Maladie des Travailleurs Salariés, Direction Générale de la Santé, MGEN, Institut de la Longévité, Agence Française de Sécurité Sanitaire des Produits de Santé, the Aquitaine and Bourgogne Regional Councils, Agence Nationale de la Recherche, ANR supported the COGINUT and COVADIS projects. Fondation de France and the joint French Ministry of Research/INSERM "Cohortes et collections de données biologiques" programme. Lille Génopôle received an unconditional grant from Eisai. The Three-city biological bank was developed and maintained by the laboratory for genomic analysis LAG-BRC - Institut Pasteur de Lille. This work was further supported by the CoSTREAM project (http://www.costream.eu/) and funding from the European Union's Horizon 2020 research and innovation program under grant agreement 667375.

Belgium samples: Research at the Antwerp site is funded in part by the Belgian Science Policy Office Interuniversity Attraction Poles program, the Belgian Alzheimer Research Foundation, the Flemish government-initiated Flanders Impulse Program on Networks for Dementia Research (VIND) and the Methusalem excellence program, the Research Foundation Flanders (FWO), and the University of Antwerp Research Fund, Belgium. The Antwerp site authors thank the personnel of the VIB Neuromics Support Facility, the Biobank of the Institute Born-Bunge and neurology departments at the contributing hospitals. The authors acknowledge the members of the BELNEU consortium for their contributions to the clinical and pathological characterization of Belgium patients and the personnel of the Diagnostic Service Facility for the genetic testing.

Finish sample collection: Financial support for this project was provided by Academy of Finland (grant number 307866), Sigrid Jusélius Foundation and the Strategic Neuroscience Funding of the University of Eastern Finland

Swedish sample collection: Financially supported in part by the Swedish Brain Power network, the Marianne and Marcus Wallenberg Foundation, the Swedish Research Council (521-2010-3134, 2015-02926), the King Gustaf V and Queen Victoria's Foundation of Freemasons, the Regional Agreement on Medical Training and Clinical Research (ALF) between Stockholm County Council and the Karolinska Institutet, the Swedish Brain Foundation and the Swedish Alzheimer Foundation".

CHARGE. Infrastructure for the CHARGE Consortium is supported in part by National Heart, Lung, and Blood Institute grant HL105756 (Psaty) and RC2HL102419 (Boerwinkle) and the neurology working group by grants from the National Institute on Aging, R01 AG033193, U01 AG049505 and U01AG52409.

Rotterdam (RS). This study was funded by the Netherlands Organisation for Health Research and Development (ZonMW) as part of the Joint Programming for Neurological Disease (JPND)as part of the PERADES Program (Defining Genetic Polygenic, and Environmental Risk for Alzheimer's disease using multiple powerful cohorts, focused Epigenetics and Stem cell metabolomics), Project number 733051021. This work was funded also by the European Union Innovative Medicine Initiative (IMI) programme under grant agreement No. 115975 as part of the Alzheimer's Disease Apolipoprotein Pathology for Treatment Elucidation and Development (ADAPTED, https://www.imi-adapted.eu);and the European Union's Horizon 2020 research and innovation programme as part of the Common mechanisms and pathways in Stroke and Alzheimer's disease CoSTREAM project (www.costream.eu, grant agreement No. 667375). The current study is supported by the Deltaplan Dementie and Memorabel supported by ZonMW (Project number 733050814) and Alzheimer Nederland.

The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. The generation and management of GWAS genotype data for the Rotterdam Study (RS-I, RS-II, RS-III) was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands. The GWAS datasets are supported by the Netherlands Organization of Scientific Research NWO Investments (Project number 175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA), project number 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters, MSc, and Carolina Medina-Gomez, MSc, for their help in creating the GWAS database, and Karol Estrada, PhD, Yurii Aulchenko, PhD, and Carolina Medina-Gomez, MSc, for the creation and analysis of imputed data.

AGES. The AGES study has been funded by NIA contracts N01-AG-12100 and HHSN271201200022C with contributions from NEI, NIDCD, and NHLBI, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament).

Cardiovascular Health Study (CHS). This research was supported by contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, and N01HC85086 and grant U01HL080295 *and*

U01HL130114 from the National Heart, Lung, and Blood Institute (NHLBI), with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided by R01AG033193, R01AG023629, R01AG15928, and R01AG20098 and by U01AG049505 from the National Institute on Aging (NIA). The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. A full list of CHS principal investigators and institutions can be found at https://chs-nhlbi.org/. The content is solely the responsibility of the authors and does not necessarily represent the official views of the US National Institutes of Health.

Framingham Heart Study. This work was supported by the National Heart, Lung, and Blood Institute's Framingham Heart Study (contracts N01-HC-25195 and HHSN268201500001I). This study was also supported by grants from the National Institute on Aging: R01AG033193, U01AG049505, U01AG52409, R01AG054076 (S. Seshadri). S. Seshadri and A.L.D. were also supported by additional grants from the National Institute on Aging (R01AG049607, R01AG033040) and the National Institute of Neurological Disorders and Stroke (R01-NS017950, NS100605). The content is solely the responsibility of the authors and does not necessarily represent the official views of the US National Institutes of Health.

GR@ACE cohort. Fundació ACE We would like to thank patients and controls who participated in this project. Genome Resesarch @ Fundació ACE project (GR@ACE) is supported by Fundación bancaria “La Caixa”, Grifols SA, Fundació ACE and ISCIII. We also want to thank other private sponsors supporting the basic and clinical projects of our institution (Piramal AG, Laboratorios Echevarne, Araclon Biotech S.A. and Fundació ACE). We are indebted to Trinitat Port-Carbó legacy and her family for their support of Fundació ACE research programs. Fundació ACE collaborates with the Centro de Investigación Biomédica en Red sobre

Enfermedades Neurodegenerativas (CIBERNED, Spain) and is one of the participating centers of the Dementia Genetics Spanish Consortium (DEGESCO). A.R. and M.B. are receiving support from the European Union/EFPIA Innovative Medicines Initiative Joint Undertaking ADAPTED and MOPEAD projects (Grants No. 115975 and 115985 respectively). M.B. and A.R. are also supported by national grants PI13/02434, PI16/01861 and PI17/01474. Acción Estratégica en Salud integrated in the Spanish National R + D + I Plan and funded by ISCIII (Instituto de Salud Carlos III)-Subdirección General de Evaluación and the Fondo Europeo de Desarrollo Regional (FEDER- "Una manera de Hacer Europa"). Control samples and data from patients included in this study were provided in part by the National DNA Bank Carlos III (www.bancoadn.org, University of Salamanca, Spain) and Hospital Universitario Virgen de Valme (Sevilla, Spain) and they were processed following standard operating procedures with the appropriate approval of the Ethical and Scientific Committee.

GERAD/PERADES. We thank all individuals who participated in this study. Cardiff University was supported by the Wellcome Trust, Alzheimer's Society (AS; grant RF014/164), the Medical Research Council (MRC; grants G0801418/1, MR/K013041/1, MR/L023784/1), the European Joint Programme for Neurodegenerative Disease (JPND, grant MR/L501517/1), Alzheimer's Research UK (ARUK, grant ARUK-PG2014-1), Welsh Assembly Government (grant SGR544:CADR), a donation from the Moondance Charitable Foundation, and the UK Dementia Research Institute at Cardiff. Cambridge University acknowledges support from the MRC. ARUK supported sample collections at the Kings College London, the South West Dementia Bank, Universities of Cambridge, Nottingham, Manchester and Belfast. King's College London was supported by the NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at the South London and Maudsley NHS Foundation Trust and Kings College London and the MRC. Alzheimer's Research UK (ARUK) and the Big Lottery Fund provided support to Nottingham University. Ulster Garden Villages, AS, ARUK, American

Federation for Aging Research, NI R&D Office and the Royal College of Physicians/Dunhill Medical Trust provided support for Queen's University, Belfast. The University of Southampton acknowledges support from the AS. The MRC and Mercer's Institute for Research on Ageing supported the Trinity College group. DCR is a Wellcome Trust Principal Research fellow. The South West Dementia Brain Bank acknowledges support from Bristol Research into Alzheimer's and Care of the Elderly. The Charles Wolfson Charitable Trust supported the OPTIMA group. Washington University was funded by NIH grants, Barnes Jewish Foundation and the Charles and Joanne Knight Alzheimer's Research Initiative. Patient recruitment for the MRC Prion Unit/UCL Department of Neurodegenerative Disease collection was supported by the UCLH/UCL Biomed- ical Centre and their work was supported by the NIHR Queen Square Dementia BRU. LASER-AD was funded by Lundbeck SA. The Bonn group would like to thank Dr. Heike Koelsch for her scientific support. The Bonn group was funded by the German Federal Ministry of Education and Research (BMBF): Competence Network Dementia (CND) grant number 01GI0102, 01GI0711, 01GI0420. The AgeCoDe study group was supported by the German Federal Ministry for Education and Research grants 01 GI 0710, 01 GI 0712, 01 GI 0713, 01 GI 0714, 01 GI 0715, 01 GI 0716, 01 GI 0717. Genotyping of the Bonn case-control sample was funded by the German centre for Neurodegenerative Diseases (DZNE), Germany. The GERAD Consortium also used samples ascertained by the NIMH AD Genetics Initiative. HH was supported by a grant of the Katharina-Hardt-Foundation, Bad Homburg vor der Höhe, Germany. The KORA F4 studies were financed by Helmholtz Zentrum München; German Research Center for Environmental Health; BMBF; German National Genome Research Network and the Munich Center of Health Sciences. The Heinz Nixdorf Recall cohort was funded by the Heinz Nixdorf Foundation (Dr. Jur. G.Schmidt, Chairman) and BMBF. Coriell Cell Repositories is supported by NINDS and the Intramural Research Program of the National Institute on Aging. We acknowledge use of genotype data from the 1958 Birth Cohort collection, funded by the MRC and the Wellcome Trust which was genotyped by the Wellcome Trust Case

Control Consortium and the Type-1 Diabetes Genetics Consortium, sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Allergy and Infectious Diseases, National Human Genome Research Institute, National Institute of Child Health and Human Development and Juvenile Diabetes Research Foundation International.

The Bonn samples are part of the German Dementia Competance Network (DCN) and the German Research Network on Degenerative Dementia (KNDD), which are funded by the German Federal Ministry of Education and Research (grants KND: 01G10102, 01GI0420, 01GI0422, 01GI0423, 01GI0429, 01GI0431, 01GI0433, 04GI0434; grants KNDD: 01GI1007A, 01GI0710, 01GI0711, 01GI0712, 01GI0713, 01GI0714, 01GI0715, 01GI0716, 01ET1006B). Markus M Nothen is a member of the German Research Foundation (DFG) cluster of excellence ImmunoSensation. Funding for Saarland University was provided by the German Federal Ministry of Education and Research (BMBF), grant number 01GS08125 to Matthias Riemenschneider. The University of Washington was supported by grants from the National Institutes of Health (R01-NS085419 and R01-AG044546), the Alzheimer's Association (NIRG-11-200110) and the American Federation for Aging Research (Carlos Cruchaga was recipient of a New Investigator Award in Alzheimer's disease). Brigham Young University was supported by the Alzheimer's Association (MNIRG-11-205368), the BYU Gerontology Program and the National Institutes of Health (R01-AG11380, R01-AG021136, P30- S069329-01, R01-AG042611). We also acknowledge funding from the Institute of Neurology, UCL, London who were supported in part by the ARUK via an anonymous donor, and by a fellowship to Dr Guerreiro. Seripa, Urbano and Masullo's participation in the study was completely supported by Ministerodella Salute", I.R.C.C.S. Research Program, Ricerca Corrente 2015-2017, Linea n. 2 "Malattiecomplesse e terapie innovative" and by the "5 × 1000" voluntary contribution. AddNeuromed is supported by InnoMed, an Integrated Project funded by the European Union Sixth Framework programme priority FP6-2004-LIFESCIHEALTH-5, Life Sciences, Genomics

and Biotechnology for Health. We are grateful to the Wellcome Trust for awarding a Principal Research Fellowship to Rubensztein (095317/Z/11/Z). Matthias Riemenschneider was funded by the BMBF NGFN Grant 01GS08125. BN supported by FondazioneCassa di Risparmio di Pistoia e Pescia (grants 2014.0365, 2011.0264 and 2013.0347). Harald Hampel is supported by the AXA Research Fund, the “Fondation partenariale Sorbonne Université” and the “Fondation pour la Recherche sur Alzheimer”, Paris, France. Ce travail a bénéficié d'une aide de l’Etat “Investissements d’avenir” ANR-10-IAIHU-06. The research leading to these results has received funding from the program “Investissements d’avenir” ANR-10-IAIHU-06 (Agence Nationale de la Recherche-10-IA Agence Institut Hospitalo-Universitaire-6. The research leading to these results has received funding from the program “Investissementsd’ avenir” ANR-10-IAIHU-06 (Agence Nationale de la Recherche-10-IA Agence Institut Hospitalo-Universitaire-6. The Santa Lucia Foundation and the Fondazione Ca’ Granda IRCCS Ospedale Policlinico, Italy, acknowledge the Italian Ministry of Health (grant RC 10.11.12.13/A). We acknowledge Maria A Pastor (Department of Neurology, University of Navarra Medical School and Neuroimaging Laboratory, Center for Applied Medical Research, Pamplona, Spain), for providing DNA samples.

Author Contributions

ADGC. *Study design or conception:* A.C.N., A.A.-W., E.R.M., K.H.-N., A.B.K., B.N.V., G.W.B., O.V., M.Butkiewicz, W.B., Y.Song, G.D.S., M.A.P.-V. *Sample contribution:* S.S.M., P.K.C., R.B., P.M.A., M.S.A., D. Beekly, D. Blacker, R.S. Doody, T.J.F., M.P.F., B.Ghetti, R.M.H., M.I.K., M.J.K., C.K., W.K., E.B.L., R.B.L., T.J.M., R.C.P., E.M.R., J.S.R., D.R.R., M. Sano, P.S.G.-H., D.W.T., C.K.W., R.L.A., L.G.A., S.E.A., S.A., C.S.A., C.T.B., L.L.B., S. Barral, T.G.B., J.T.B., E.B., T.D.B., B.F.B., J.D.B., A.Boxer, J.R.B., J.M.B., J.D.Buxbaum, N.J.C., C. Cao, C.S.C., C.M.C., R.M.C., M.M.C., H.C.C., D.H.C., E.A.C., C.DeCarli, M.Dick, R.D., N.R.G.-R., D.A.E., K.M.F., K.B.F., D.W.F., M.R.F., S.F., T.M.F., D.R.G., M.Gearing, D.H.G., J.R.G., R.C.G., J.H.G., R.H., L.E.H., L.S.H., M.J.H., C.M.H., B.T.H., G.P.J., E.A., L.W.J., G.R.J., A. Karydas, J.A.K.,

R.K., N.W.K., J.H.K., F.M.L., J.J.L., J.B.L., A.I.L., A.P.L., K.L.L., C.G.L., D.C.M., F.M., D.C.Mash, E.M., W.C.M., S.M.M., A.N.M., A.C.M., M.M., B.L.M., C.A.M., J.W.M., J.C.M., A.J.M., S.O., J.M.O., J.E.P., H.L.P., W.P., E.P., A.P., W.W.P., H.P., J.F.Q., A.Raj, M.R., B.R., C.R., J.M.R., E.D.R., E.R., H.J.R., R.N.R., M.A.S., A.J.S., M.L.C., J. Vance, J.A.S., L.S.S., S.Slifer, W.W.S., A.G.S., J.A.Sonnen, S. Spina, R.A.S., R.H.S., R.E.T., J.Q.T., J.C.T., V.M.V.D., L.J.V.E., H.V.V., J.P.V., S.W., K.A.W.-B., K.C.W., J.Williamson, T.S.W., R.L.W., C.B.W., C.-E.Y., L.Y., D.B., P.L.D.J., S.McDonough, C.Cruchaga, A.M.G., N.E.-T., S.G.Y., D.W.D., H.H., L.A.F., J.Haines, R.Mayeux, L.-S.W., G.D.S., M.A.P.-V. *Data generation:* B.W.K., K.H.-N., A.B.K., O.V., L.Q., Y.Z., J.Malamon, B.Dombroski, P.W., L.B.C., M.A., M.Tang, J.R.G., L.-S.W. *Analysis:* B.W.K., A.C.N., A.A.-W., E.R.M., K.H.-N., A.B.K., M.Tang, B.N.V., G.W.B., O.V., M.Butkiewicz, W.B., Y.S., G.D.S., M.A.P.-V. *Manuscript preparation:* B.W.K., G.D.S., M.A.P.-V. *Study supervision/management:* B.W.K., L.A.F., J.Haines, R.Mayeux, L.-S.W., G.D.S., M.A.P.-V.

EADI. *Study design or conception:* P.A., J.-C.L. *Sample contribution:* K.S., M.Hiltunen, J.E., M.D.Z., I.M., F.S.-G., M.C.D.N., D.Wallon, S.E., R.V., P.D.D., A.Squassina, E.R.-R., C.M.-F., Y.A.B., H.T., V.Giedraitis, L.Kilander, R.Brundin, L.C., S.Helisalmi, A.M.K., A.Haapasalo, V.S., V.Frisardi, V.Deramecourt, N.F., O.H., C.Dufouil, A.Brice, K.R., B.D., H.Soininen, L.Fratiglioni, L.K., F.Panza, D.H., P.C., F.S., P.B., L.Lannfelt, F.P., M.Ingelsson, C.G., P.S.-J., A.L., J.Clarimon, C.Berr, S.D., J.-F.D., A.Pilotto, M.J.B., P.Bosco, E.C., G.N., D.C., C.V.B.,J.-F.Deleuze, P.A., J.-C.L. *Data generation:* R.O., J.-G.G., M.-L.M., D.Bacq, F.G., B.F., S.Meslage *Analysis:* B.G.-B., V.D., A.Boland, C.Bellenguez *Manuscript preparation:* B.G.-B., P.A., J.-C.L. *Study supervision/management:* P.A., J.-C.L.

GERAD/PERADES. *Study design or conception:* R.Sims, M.C.O., M.J.O., A.R., P.A.H., J.W. *Sample contribution:* R.Raybould, T.Morgan, P.Hoffman, D.Harold, O.P., N.D., N.C.F., J.T.H., Y.P., M.Daniilidou, J.U., D.Galimberti, E.Scarpini, J.Kornhuber, S.P., M.Mayhaus, W.G., A.M.H., S.Lovestone, R.Sussams, C.Holmes, W.M., A.Kawalia, S.Moebus, J.Turton, J.Lord, I.K., A.L., B.L., M.Gill, S.O.-C., C.Caltagirone, C.Cupidi, R.G.M., R.Cecchetti, M.T., D.Craig, D.A., A.G., M.K., O.G., D.Makrina, H.Hampel, D.C.R., L.F., B.G., J.J., P.Passmore, J.M.S., J.D.W., M.K.L., P.Proitsi, J.Powell, J.S.K.K., M.Mancuso, U.B., A.Q., G.Livingston, N.J.B., J.Hardy, J.B., R.Guerreiro, E.F., C.Masullo, G.B., L.M., A.H., M.Scherer, M.Reimenschneider, R.Heun, M.Leber, I.H., I.G., M.Hull, J.M., K.Mayo, T.F., D.Drichel, T.D.C., P.Hollingworth, R.Marshall, A.Meggy, G.M., G.L., D.G., G.R., F.J., B.V., E.V., K.-H.J., M.Dichgans, D.Mann, S.P.-B., N.K.,

H.W., K.M., K.Brown, C.Medway, M.M.N., N.M.H., A.Daniele, A.Bayer, J.G., H.V.D.B., C.Brayne, S.R.-H., A.A.-C., C.E.S., J.Wiltfang, V.A., A.B.S., J.C., S.M., M.Rossor, N.R., B.N., S.Sorbi, E.S., G.S., R.C., A.D.S., D.W., G.W., A.C.B., M.G., Y.B.-S., P.M., P.P., V.B., N.W., P.D., R.G., P.G.K., S.L., C.C., J.T., R.Munger, A.R., J.W. *Data generation:* R.Sims, R.Raybould, T.Morgan, P.Hoffman, D.Harold, A.Gerrish, N.D., P.Hollingworth, R.Marshall, A.Meggy, A.R., J.W. *Analysis:* R.Sims, M.V., A.F., N.Badarinarayan, D.Harold, G.M., G.L., D.G., V.E.-P., A.R., J.W. *Manuscript preparation:* R.Sims, T.D.C., P.A.H., J.W. *Study supervision/management:* R.Sims, L.J., V.E.-P., A.R., P.A.H., J.W.

CHARGE. *Study design or conception:* A.L.D., C.M.V.D., S.S. *Sample contribution:* J.C.B., A.Ruiz, I.D.R., L.M.R., I.Q., A.C., A.L.F., G.E., J.J.H., A.O., M.E.G., H.L., H.Comic, G.Roschupkin, S.Li, I.Hernández, Q.Y., A.S.B., L.T., T.H.M., WT.L., F.R., E.Boerwinkle, J.I.R., A.G.U., S.M.-G., O.L.L., M.B., M.F., N.A., L.J.L., M.A.I., H.S., R.S., V.G., B.M.P. *Data generation:* J.C.B., J.Jakobsdottir, A.Ruiz, A.V.S., X.J., S.-H.C., H.H.A., J.A.B., T.A., E.H., C.Sarnowski, D.V., L.A.C. *Analysis:* J.C.B., S.J.v.d.L., V.C., J.Jakobsdottir, Y.C., Y.Saba, S.Ahmad, A.Ruiz, A.V.S., C.C.W., C.M.V.D., S.S. *Manuscript preparation:* S.J.v.d.L., A.Ruiz, B.M.P., C.M.V.D., S.S. *Study supervision/management:* C.M.V.D., S.S.

Competing Interests statement

D. Blacker is a consultant for Biogen, Inc. R.C.P. is a consultant for Roche, Inc., Merck, Inc., Genentech, Inc., Biogen, Inc., and Eli Lilly. A.R.W. is a former employee and stockholder of Pfizer, Inc., and a current employee of the Perelman School of Medicine at the University of Pennsylvania Orphan Disease Center in partnership with the Loulou. A.M.G. is a member of the scientific advisory board for Denali Therapeutics. N.E.-T. is a consultant for Cytox. J. Hardy holds a collaborative grant with Cytox cofunded by the Department of Business (Biz). F.J. acts as a consultant for Novartis, Eli Lilly, Nutricia, MSD, Roche, and Piramal. Neither J. Morris nor his family own stock or have equity interest (outside of mutual funds or other externally directed accounts) in any pharmaceutical or biotechnology company. J. Morris is currently participating in clinical trials of antidementia drugs from Eli Lilly and Company, Biogen, and Janssen. J. Morris serves as a consultant for Lilly USA. He receives research

support from Eli Lilly/Avid Radiopharmaceuticals and is funded by NIH grants P50AG005681, P01AG003991, P01AG026276, and UF01AG032438.

References

1. Association, A. *et al.* 2014 Alzheimer's disease facts and figures. *Alzheimers Dement* **8,** e47–e92 (2013).
2. Adams, P. M. *et al.* Assessment of the genetic variance of late-onset Alzheimer's disease. *Neurobiol. Aging* **41,** 1–8 (2016).
3. Lambert, J. C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45,** 1452–8 (2013).
4. Gatz, M. *et al.* Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry* **63,** 168–174 (2006).
5. Naj, A. C. *et al.* Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet* **43,** 436–441 (2011).
6. Seshadri, S. *et al.* Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA* **303,** 1832–1840 (2010).
7. Hollingworth, P. *et al.* Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet* **43,** 429–435 (2011).
8. Jonsson, T. *et al.* Variant of TREM2 associated with the risk of Alzheimer's disease. *N. Engl. J. Med.* **368,** 107–16 (2013).
9. Guerreiro, R. *et al.* TREM2 variants in Alzheimer's disease. *N. Engl. J. Med.* **368,** 117–27 (2013).
10. Jun, G. *et al.* Meta-analysis confirms CR1, CLU, and PICALM as alzheimer disease risk loci and reveals interactions with APOE genotypes. *Arch Neurol* **67,** 1473–1484 (2010).
11. Harold, D. *et al.* Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet* **41,** 1088–1093 (2009).
12. Lambert, J. C. *et al.* Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet* **41,** 1094–1099 (2009).
13. Zheng, J. *et al.* LD Hub: A centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33,** 051094 (2017).
14. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47,** 291–295 (2015).
15. Sims, R. C. *et al.* Novel rare coding variants in PLCG2, ABI3 and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. *Nat. Genet.* 1373–1387 (2017). doi:10.1038/ng.3916

16. Vassar, R. ADAM10 prodomain mutations cause late-onset Alzheimer's disease: not just the latest FAD. *Neuron* **80,** 250–253 (2013).
17. Kim, M. *et al.* Potential late-onset Alzheimer's disease-associated mutations in the ADAM10 gene attenuate {alpha}-secretase activity. *Hum Mol Genet* **18,** 3987–3996 (2009).
18. Kehoe, P. G. *et al.* Variation in DCP1, encoding ACE, is associated with susceptibility to Alzheimer disease. *Nat. Genet.* **21,** 71–72 (1999).
19. Meng, Y. *et al.* Association of polymorphisms in the Angiotensin-converting enzyme gene with Alzheimer disease in an Israeli Arab community. *Am J Hum Genet* **78,** 871–877 (2006).
20. Lehmann, D. J. *et al.* Large meta-analysis establishes the ACE insertion-deletion polymorphism as a marker of Alzheimer's disease. *Am. J. Epidemiol.* **162,** 305–17 (2005).
21. Wang, X.-B. *et al.* Angiotensin-Converting Enzyme Insertion/Deletion Polymorphism Is Not a Major Determining Factor in the Development of Sporadic Alzheimer Disease: Evidence from an Updated Meta-Analysis. *PLoS One* **9,** e111406 (2014).
22. Cai, G. *et al.* Evidence against a role for rare ADAM10 mutations in sporadic Alzheimer disease. *Neurobiol Aging* **33,** 416–417 e3 (2012).
23. Belbin, O. *et al.* A multi-center study of ACE and the risk of late-onset Alzheimer's disease. *J Alzheimers Dis* **24,** 587–597 (2011).
24. Wakutani, Y. *et al.* Genetic analysis of vascular factors in Alzheimer's disease. *Ann. N. Y. Acad. Sci.* **977,** 232–8 (2002).
25. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet* **44,** 369–75, S1-3 (2012).
26. Chang, J.-Y. & Chang, N.-S. WWOX dysfunction induces sequential aggregation of TRAPPC6AΔ, TIAF1, tau and amyloid β, and causes apoptosis. *Cell Death Discov.* **1,** 15003 (2015).
27. Sze, C. I. *et al.* Down-regulation of WW domain-containing oxidoreductase induces Tau phosphorylation in vitro: A potential role in Alzheimer's disease. *J. Biol. Chem.* **279,** 30498–30506 (2004).
28. Zhang, B. *et al.* Integrated Systems Approach Identifies Genetic Nodes and Networks in Late-Onset Alzheimer's Disease. *Cell* **153,** 707–720 (2013).
29. Bai, Z. *et al.* AlzBase: an Integrative Database for Gene Dysregulation in Alzheimer's

Disease. *Mol. Neurobiol.* **53,** 310–319 (2016).

30. Zhang, Y. *et al.* An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* **34,** 11929–11947 (2014).
31. Zhang, Y. *et al.* Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. *Neuron* **89,** 37–53 (2016).
32. Olah, M. *et al.* A transcriptomic atlas of aged human microglia. *Nat. Commun.* **9,** 1–8 (2018).
33. Corder, E. H. *et al.* Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nat. Genet.* **7,** 180–184 (1994).
34. Kim, J., Basak, J. M. & Holtzman, D. M. The role of apolipoprotein E in Alzheimer's disease. *Neuron* **63,** 287–303 (2009).
35. Steinberg, S. *et al.* Loss-of-function variants in ABCA7 confer risk of Alzheimer's disease. *Nat. Genet.* 26–29 (2015). doi:10.1038/ng.3246
36. Vasquez, J. B., Fardo, D. W. & Estus, S. ABCA7 expression is associated with Alzheimer's disease polymorphism and disease status. *Neurosci Lett* **556,** 58–62 (2013).
37. De Roeck, A. *et al.* An intronic VNTR affects splicing of ABCA7 and increases risk of Alzheimer's disease. *Acta Neuropathol.* **135,** 827–837 (2018).
38. De Roeck, A. *et al.* Deleterious ABCA7 mutations and transcript rescue mechanisms in early onset Alzheimer's disease. *Acta Neuropathol.* **134,** 475–487 (2017).
39. Chapuis, J. *et al.* Increased expression of BIN1 mediates Alzheimer genetic risk by modulating tau pathology. *Mol Psychiatry* **18,** 1225–34 (2013).
40. Rogaeva, E. *et al.* The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. *Nat Genet* **39,** 168–177 (2007).
41. Vardarajan, B. N. *et al.* Coding mutations in *SORL 1* and Alzheimer disease. *Ann. Neurol.* **77,** 215–227 (2015).
42. Suh, J. *et al.* ADAM10 missense mutations potentiate beta-amyloid accumulation by impairing prodomain chaperone function. *Neuron* **80,** 385–401 (2013).
43. Huang, K. *et al.* A common haplotype lowers PU.1 expression in myeloid cells and delays onset of Alzheimer's disease. *Nat. Neurosci.* **20,** 1052–1061 (2017).
44. Brouwers, N. *et al.* Alzheimer risk associated with a copy number variation in the complement receptor 1 increasing C3b/C4b binding sites. *Mol Psychiatry* **17,** 223–233 (2012).
45. Flister, M. J. *et al.* Identifying multiple causative genes at a single GWAS locus. *Genome*

Res. **23,** 1996–2002 (2013).

46. Farh, K. K.-H. *et al.* Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* (2014). doi:10.1038/nature13835
47. Bis, J. C. *et al.* Whole exome sequencing study identifies novel rare and common Alzheimer's-associated variants involved in immune response and transcriptional regulation. *Mol. Psychiatry* Accepted May (2018).
48. Vardarajan, B. N. *et al.* Coding mutations in SORL1 and Alzheimer disease. *Ann. Neurol.* **77,** 215–227 (2015).
49. Verheijen, J. *et al.* A comprehensive study of the genetic impact of rare variants in SORL1 in European early-onset Alzheimer's disease. *Acta Neuropathol.* (2016). doi:10.1007/s00401-016-1566-9
50. Bellenguez, C. *et al.* Contribution to Alzheimer's disease risk of rare variants in TREM2, SORL1, and ABCA7 in 1779 cases and 1273 controls. *Neurobiol. Aging* **59,** 220.e1-220.e9 (2017).
51. Kunkle, B. W. *et al.* Targeted sequencing of ABCA7 identifies splicing, stop-gain and intronic risk variants for Alzheimer disease. *Neurosci. Lett.* **649,** 124–129 (2017).
52. May, P. *et al.* Rare ABCA7 variants in 2 German families with Alzheimer disease. *Neurol. Genet.* **4,** e224 (2018).
53. Guennec, K. Le *et al.* ABCA7 rare variants and Alzheimer disease risk. *Neurology* **86,** 1–4 (2016).
54. Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22,** 1790–7 (2012).
55. Ward, L. D. & Kellis, M. HaploReg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* **40,** D930-4 (2012).
56. Zerbino, D. R., Wilder, S. P., Johnson, N., Juettemann, T. & Flicek, P. R. The Ensembl Regulatory Build. *Genome Biol.* **16,** 56 (2015).
57. Huang, D. *et al.* GWAS4D: multidimensional analysis of context-specific regulatory variant for human complex diseases and traits. *Nucleic Acids Res.* **46,** W114–W120 (2018).
58. Gjoneska, E. *et al.* Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. *Nature* **518,** 365–369 (2015).
59. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput. Biol.* **11,** 1–19 (2015).

60. Stefanis, L. alpha-Synuclein in Parkinson's disease. *Cold Spring Harb. Perspect. Med.* **2,** 1–23 (2012).
61. Takeda, A. *et al.* C-terminal alpha-synuclein immunoreactivity in structures other than Lewy bodies in neurodegenerative disorders. *Acta Neuropathol.* **99,** 296–304 (2000).
62. Campion, D., Pottier, C., Nicolas, G., Le Guennec, K. & Rovelet-Lecrux, A. Alzheimer disease: modeling an Aβ-centered biological network. *Mol. Psychiatry* 1–11 (2016). doi:10.1038/mp.2016.38
63. Yeh, F. L., Wang, Y., Tom, I., Gonzalez, L. C. & Sheng, M. TREM2 Binds to Apolipoproteins, Including APOE and CLU/APOJ, and Thereby Facilitates Uptake of Amyloid-Beta by Microglia. *Neuron* **91,** 328–340 (2016).
64. Fritsche, L. G. *et al.* A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. *Nat. Genet.* (2015). doi:10.1038/ng.3448
65. Haass, C., Kaether, C., Thinakaran, G. & Sisodia, S. Trafficking and proteolytic processing of APP. *Cold Spring Harb Perspect Med* **2,** a006270 (2012).
66. Kleinberger, G. *et al.* TREM2 mutations implicated in neurodegeneration impair cell surface transport and phagocytosis. *Sci. Transl. Med.* **6,** 243ra86-243ra86 (2014).
67. Postina, R. *et al.* A disintegrin-metalloproteinase prevents amyloid plaque formation and hippocampal defects in an Alzheimer disease mouse model. *J. Clin. Invest.* **113,** 1456–1464 (2004).
68. Hinney, A. *et al.* Genetic variation at the CELF1 (CUGBP, elav-like family member 1 gene) locus is genome-wide associated with Alzheimer's disease and obesity. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **165B,** 283–93 (2014).
69. Speliotes, E. K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42,** 937–48 (2010).
70. Kurabayashi, N., Nguyen, M. D. & Sanada, K. The G protein-coupled receptor GPRC5B contributes to neurogenesis in the developing mouse neocortex. *Development* **140,** 4335–4346 (2013).
71. Cool, B. H. *et al.* A flanking gene problem leads to the discovery of a Gprc5b splice variant predominantly expressed in C57BL/6J mouse brain and in maturing neurons. *PLoS One* **5,** (2010).
72. Kim, Y.-J., Sano, T., Nabetani, T., Asano, Y. & Hirabayashi, Y. GPRC5B Activates Obesity-Associated Inflammatory Signaling in Adipocytes. *Sci. Signal.* **5,** ra85-ra85 (2012).

73. Bhat, K. *et al.* The 19S proteasome ATPase Sug1 plays a critical role in regulating MHC class II transcription. *Mol. Immunol.* **45,** 2214–24 (2008).
74. Inostroza-Nieves, Y., Venkatraman, P. & Zavala-Ruiz, Z. Role of Sug1, a 19S proteasome ATPase, in the transcription of MHC I and the atypical MHC II molecules, HLA-DM and HLA-DO. *Immunol. Lett.* **147,** 67–74 (2012).
75. Kim, K., Duramad, O., Qin, X. F. & Su, B. MEKK3 is essential for lipopolysaccharide-induced interleukin-6 and granulocyte-macrophage colony-stimulating factor production in macrophages. *Immunology* **120,** 242–250 (2007).
76. Yamazaki, K. *et al.* Two mechanistically and temporally distinct NF-κB activation pathways in IL-1 signaling. *Sci. Signal.* **2,** 1–12 (2009).
77. Farrer, L. A. *et al.* Association Between Angiotensin-Converting Enzyme and Alzheimer Disease. *New Engl. J. Med.* **57,** 210–14 (2000).
78. Miners, J. S. *et al.* Angiotensin-converting enzyme levels and activity in Alzheimer's disease: Differences in brain and CSF ACE and association with ACE1 genotypes. *Am. J. Transl. Res.* **1,** 163–177 (2009).
79. Jochemsen, H. M. *et al.* The association of angiotensin-converting enzyme with biomarkers for Alzheimer's disease. *Alzheimer's Res. Ther.* **6,** 1–10 (2014).
80. Kauwe, J. S. K. *et al.* Genome-Wide Association Study of CSF Levels of 59 Alzheimer's Disease Candidate Proteins: Significant Associations with Proteins Involved in Amyloid Processing and Inflammation. *PLoS Genet.* **10,** e1004758 (2014).
81. Baranello, R. J. *et al.* Amyloid-beta protein clearance and degradation (ABCD) pathways and their role in Alzheimer's disease. *Curr. Alzheimer Res.* **12,** 32–46 (2015).
82. Kehoe, P. G. The Coming of Age of the Angiotensin Hypothesis in Alzheimer's Disease: Progress Toward Disease Prevention and Treatment? *J. Alzheimer 's Dis.* **62,** In Press (2018).
83. Kehoe, P. G. *et al.* The Rationale and Design of the Reducing Pathology in Alzheimer's Disease through Angiotensin TaRgeting (RADAR) Trial. *J. Alzheimer's Dis.* **61,** 803–814 (2017).
84. Miguel, R. F., Pollak, A. & Lubec, G. Metalloproteinase ADAMTS-1 but not ADAMTS-5 is manifold overexpressed in neurodegenerative disorders as Down syndrome, Alzheimer's and Pick's disease. *Brain Res. Mol. Brain Res.* **133,** 1–5 (2005).
85. Suttkus, A. *et al.* Aggrecan, link protein and tenascin-R are essential components of the perineuronal net to protect neurons against iron-induced oxidative stress. *Cell Death Dis.* **5,** e1119 (2014).

86. Végh, M. J. *et al.* Reducing hippocampal extracellular matrix reverses early memory deficits in a mouse model of Alzheimer's disease. *Acta Neuropathol. Commun.* **2,** 76 (2014).
87. Morawski, M., Filippov, M., Tzinia, A., Tsilibary, E. & Vargova, L. ECM in brain aging and dementia. *Prog. Brain Res.* **214,** 207–27 (2014).
88. Wilcock, D. M. Neuroinflammation in the aging down syndrome brain; lessons from Alzheimer's disease. *Curr Gerontol Geriatr Res* **2012,** 170276 (2012).
89. Wang, K. *et al.* A genome-wide association study on obesity and obesity-related traits. *PLoS One* **6,** 3–8 (2011).
90. Kang, K. *et al.* Interferon-γ Represses M2 Gene Expression in Human Macrophages by Disassembling Enhancers Bound by the Transcription Factor MAF. *Immunity* **47,** 235–250.e4 (2017).
91. Cao, S., Liu, J., Song, L. & Ma, X. The Protooncogene c-Maf Is an Essential Transcription Factor for IL-10 Gene Expression in Macrophages. *J. Immunol.* **174,** 3484–3492 (2005).
92. Lee, J. C. *et al.* WW-Domain-Containing Oxidoreductase Is Associated with Low Plasma HDL-C Levels. *Am. J. Hum. Genet.* **83,** 180–192 (2008).
93. Saez, M. E. *et al.* WWOX gene is associated with HDL cholesterol and triglyceride levels. *BMC Med. Genet.* **11,** 148 (2010).
94. Chang, H. T. *et al.* WW domain-containing oxidoreductase in neuronal injury and neurological diseases. *Oncotarget* **5,** 11792–11799 (2014).
95. Lee, M. H. *et al.* Zfra restores memory deficits in Alzheimer's disease triple-transgenic mice by blocking aggregation of TRAPPC6AΔ, SH3GLB2, tau, and amyloid β, and inflammatory NF-κB activation. *Alzheimer's Dement. Transl. Res. Clin. Interv.* **3,** 189–204 (2017).
96. Dourlen, P. *et al.* Functional screening of Alzheimer risk loci identifies PTK2B as an in vivo modulator and early marker of Tau pathology. *Mol. Psychiatry* **22,** 874–883 (2017).
97. Chapuis, J. *et al.* Genome-wide, high-content siRNA screening identifies the Alzheimer's genetic risk factor FERMT2 as a major modulator of APP metabolism. *Acta Neuropathol.* **133,** 955–966 (2017).
98. Shulman, J. M. *et al.* Functional screening in Drosophila identifies Alzheimer's disease susceptibility genes and implicates tau-mediated mechanisms. *Hum. Mol. Genet.* **23,** 870–877 (2014).
99. Zhao, Z. *et al.* Central role for PICALM in amyloid-β blood-brain barrier transcytosis and clearance. *Nat. Neurosci.* **18,** (2015).

100. Aguet, F. *et al.* Genetic effects on gene expression across human tissues. *Nature* **550,** 204–213 (2017).
101. Miller, J. A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508,** 199–206 (2014).
102. Knickmeyer, R. C. & Ross, M. E. Imaging and rare APOE alleles. *Neurology* 10.1212/WNL.0000000000002956 (2016). doi:10.1212/WNL.0000000000002956
103. Douaud, G. *et al.* A common brain network links development , aging , and vulnerability to disease. *Proc. Natl. Acad. Sci.* **111,** 201410378 (2014).
104. Steele, N. Z. *et al.* Fine-mapping of the human leukocyte antigen locus as a risk factor for Alzheimer disease: A case-control study. *PLoS Med.* **14,** 1–25 (2017).
105. Fekih Mrissa, N. *et al.* Association of HLA-DR-DQ polymorphisms with diabetes in Tunisian patients. *Transfus. Apher. Sci.* **49,** 200–204 (2013).
106. Pugliese, A. *et al.* HLA-DRB1∗15:01-DQA1∗01:02-DQB1∗06:02 haplotype protects autoantibody-positive relatives from type 1 diabetes throughout the stages of disease progression. *Diabetes* **65,** 1109–1119 (2016).
107. Patsopoulos, N. a *et al.* Fine-mapping the genetic association of the major histocompatibility complex in multiple sclerosis: HLA and non-HLA effects. *PLoS Genet.* **9,** e1003926 (2013).
108. Schmidt, H., Williamson, D. & Ashley-Koch, A. HLA-DR15 haplotype and multiple sclerosis: A HuGE review. *Am. J. Epidemiol.* **165,** 1097–1109 (2007).
109. Karnes, J. H. *et al.* Phenome-wide scanning identifies multiple diseases and disease severity phenotypes associated with HLA variants. *Sci. Transl. Med.* **9,** 1–14 (2017).
110. Wissemann, W. T. *et al.* Association of Parkinson disease with structural and regulatory variants in the HLA region. *Am J Hum Genet* **93,** 984–993 (2013).
111. Misra, M. K., Damotte, V. & Hollenbach, J. A. The immunogenetics of neurological disease. *Immunology* **153,** 399–414 (2018).
112. Tan, Z. S. Thyroid Function and the Risk of Alzheimer Disease: The Framingham Study. *Arch. Intern. Med.* **168,** 1514 (2008).
113. Dendrou, C. A., Petersen, J., Rossjohn, J. & Fugger, L. HLA variation and disease. *Nat. Rev. Immunol.* **18,** 325–339 (2018).
114. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47,** 1236–1241 (2015).
115. Marioni, R. E. *et al.* GWAS on family history of Alzheimer's disease. *Transl. Psychiatry* **8,** 99 (2018).

116. Stern, Y. Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurol.* **11,** 1006–1012 (2012).
117. Cadar, D. *et al.* Individual and Area-Based Socioeconomic Factors Associated With Dementia Incidence in England: Evidence From a 12-Year Follow-up in the English Longitudinal Study of Ageing. *JAMA psychiatry* **75,** 723–732 (2018).
118. Marden, J. R., Tchetgen Tchetgen, E. J., Kawachi, I. & Glymour, M. M. Contribution of Socioeconomic Status at 3 Life-Course Periods to Late-Life Memory Function and Decline: Early and Late Predictors of Dementia Risk. *Am. J. Epidemiol.* **186,** 805–814 (2017).
119. Østergaard, S. D. S. D. *et al.* Associations between Potentially Modifiable Risk Factors and Alzheimer Disease: A Mendelian Randomization Study. *PLOS Med.* **12,** e1001841 (2015).
120. Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9,** (2018).
121. Baumgart, M. *et al.* Summary of the evidence on modifiable risk factors for cognitive decline and dementia: A population-based perspective. *Alzheimers. Dement.* **11,** 1–9 (2015).
122. Larsson, S. C., Traylor, M., Burgess, S. & Markus, H. S. Genetically-Predicted Adult Height and Alzheimer's Disease. *J. Alzheimer's Dis.* **60,** 691–698 (2017).
123. Helzner, E. P. *et al.* Contribution of vascular risk factors to the progression in Alzheimer disease. *Arch. Neurol.* **66,** 343–8 (2009).
124. Reitz, C. *et al.* Association of higher levels of high-density lipoprotein cholesterol in elderly individuals and lower risk of late-onset Alzheimer disease. *Arch. Neurol.* **67,** 1491–7 (2010).
125. Mukherjee, S. *et al.* Genetically predicted body mass index and Alzheimer's disease-related phenotypes in three large samples: Mendelian randomization analyses. *Alzheimer's Dement.* **11,** (2015).
126. Arvanitakis, Z. *et al.* Late-life blood pressure association with cerebrovascular and Alzheimer disease pathology. *Neurology* 10.1212/WNL.0000000000005951 (2018). doi:10.1212/WNL.0000000000005951
127. Kuźma, E. *et al.* Which Risk Factors Causally Influence Dementia? A Systematic Review of Mendelian Randomization Studies. *J. Alzheimers. Dis.* **36,** 215–221 (2018).
128. Murray, M. E. *et al.* Clinicopathologic and 11C-Pittsburgh compound B implications of Thal amyloid phase across the Alzheimer's disease spectrum. *Brain* 1–12 (2015).

doi:10.1093/brain/awv050

129. Shi, Y. *et al.* ApoE4 markedly exacerbates tau-mediated neurodegeneration in a mouse model of tauopathy. *Nature* **549,** 523–527 (2017).

130. Brier, M. R. M. R. *et al.* Tau and Aβ imaging, CSF measures, and cognition in Alzheimer's disease. *Sci. Transl. Med.* **8,** 338ra66 (2016).

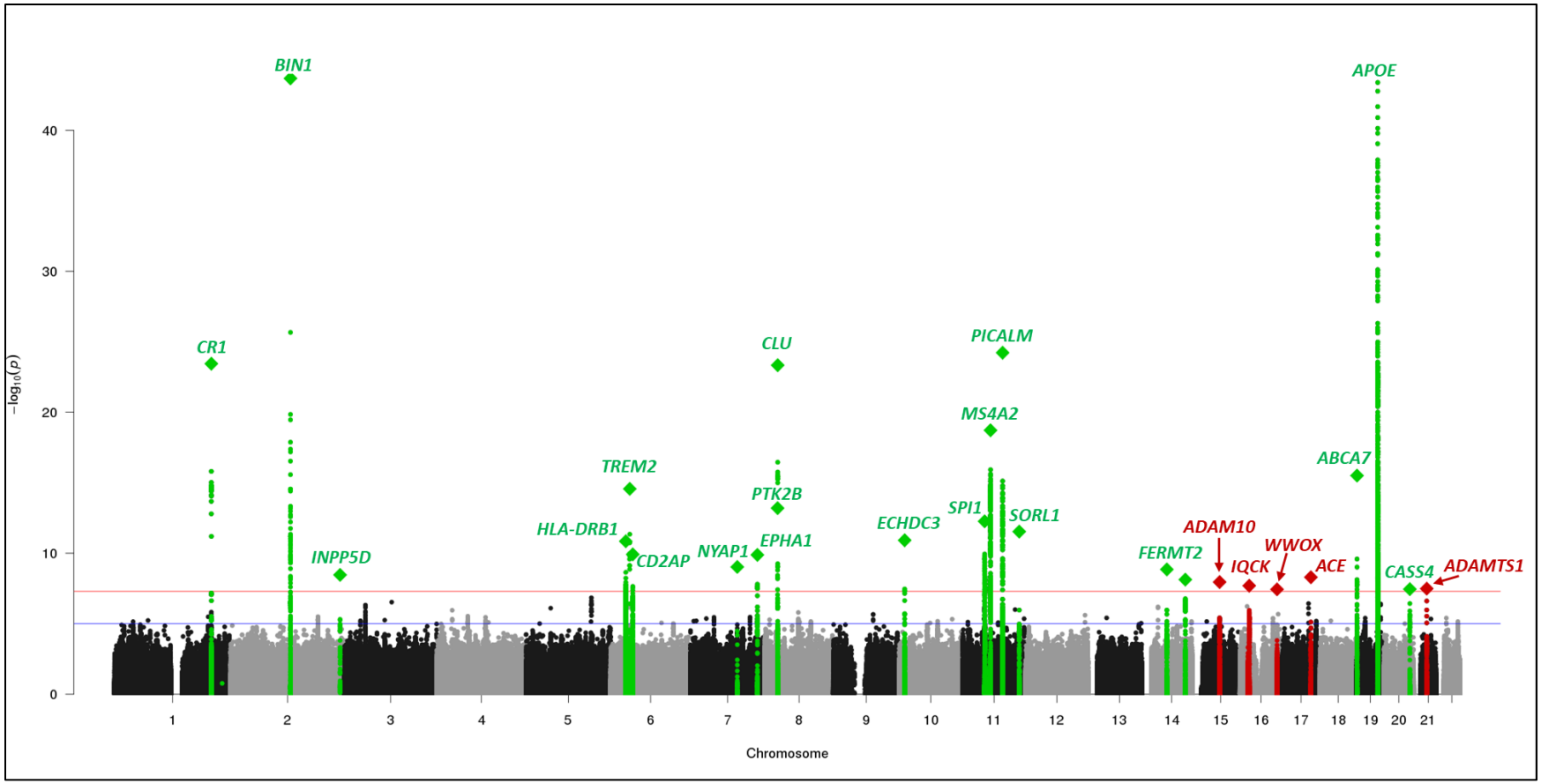

Figure 1. Manhattan plot of meta-analysis of Stage 1, 2 and 3 results for genome-wide association with Alzheimer's disease. The threshold for genome-wide significance ($P < 5 \times 10^{-8}$) is indicated by the red line, while the blue line represents the suggestive threshold ($P < 1$ x 10-5). Loci previously identified by the Lambert et al. 2013 IGAP GWAS are shown in green, and newly associated loci are shown in red. Loci are named for the closet gene to the sentinel variant for each locus. Diamonds represent variants with the smallest *P* values for each genome-wide locus.

Meta-analysis of genetic association with diagnosed Alzheimer's disease identifies novel risk loci and implicates Abeta, Tau, immunity and lipid processing - Tables

Table 1. Summary of discovery stage 1, stage 2 and overall meta-analyses results for identified loci reaching genome-wide significance after stages 1 and 2.

						Stage 1 Discovery (n=63,926)		Stage 2 (n=18,845)		Overall Stage 1 + Stage 2 (n=82,771)		
Variant[a]	Chr.	Position[b]	Closest gene[c]	Major/ minor alleles	MAF[d]	OR (95% CI)[e]	*P*	OR (95% CI)[e]	*P*	OR (95% CI)[e]	Meta *P*	I^2 (%), *P*[f]
Previous genome-wide significant loci still reaching significance												
rs4844610	1	207802552	*CR1*	C/A	0.187	1.16 (1.12-1.20)	8.2×10^{-16}	1.20 (1.13-1.27)	3.8×10^{-10}	1.17 (1.13-1.21)	3.6×10^{-24}	$0, 8 \times 10^{-1}$
rs6733839	2	127892810	*BIN1*	C/T	0.407	1.18 (1.15-1.22)	4.0×10^{-28}	1.23 (1.18-1.29)	2.0×10^{-18}	1.20 (1.17-1.23)	2.1×10^{-44}	$15, 2 \times 10^{-1}$
rs10933431	2	233981912	*INPP5D*	C/G	0.223	0.90 (0.87-0.94)	2.6×10^{-7}	0.92 (0.87-0.97)	3.2×10^{-3}	0.91 (0.88-0.94)	3.4×10^{-9}	$0, 8 \times 10^{-1}$
rs9271058	6	32575406	*HLA-DRB1*	T/A	0.270	1.10 (1.06-1.14)	5.1×10^{-8}	1.11 (1.06-1.17)	5.7×10^{-5}	1.10 (1.07-1.13)	1.4×10^{-11}	$10, 3 \times 10^{-1}$
rs75932628	6	41129252	*TREM2*	C/T	0.008	2.01 (1.65-2.44)	2.9×10^{-12}	2.50 (1.56-4.00)	1.5×10^{-4}	2.08 (1.73-2.49)	2.7×10^{-15}	$0, 6 \times 10^{-1}$
rs9473117	6	47431284	*CD2AP*	A/C	0.280	1.09 (1.05-1.12)	2.3×10^{-7}	1.11 (1.05-1.16)	1.0×10^{-4}	1.09 (1.06-1.12)	1.2×10^{-10}	$0, 6 \times 10^{-1}$
rs12539172	7	100091795	*NYAP1*[g]	C/T	0.303	0.93 (0.91-0.96)	2.1×10^{-5}	0.89 (0.84-0.93)	2.1×10^{-6}	0.92 (0.90-0.95)	9.3×10^{-10}	$0, 8 \times 10^{-1}$
rs10808026	7	143099133	*EPHA1*	C/A	0.199	0.90 (0.87-0.94)	3.1×10^{-8}	0.91 (0.86-0.96)	1.1×10^{-3}	0.90 (0.88-0.93)	1.3×10^{-10}	$0, 5 \times 10^{-1}$
rs73223431	8	27219987	*PTK2B*	C/T	0.367	1.10 (1.07-1.13)	8.3×10^{-10}	1.11 (1.06-1.16)	1.5×10^{-5}	1.10 (1.07-1.13)	6.3×10^{-14}	$0, 6 \times 10^{-1}$
rs9331896	8	27467686	*CLU*	T/C	0.387	0.88 (0.85-0.91)	3.6×10^{-16}	0.87 (0.83-0.91)	1.7×10^{-9}	0.88 (0.85-0.90)	4.6×10^{-24}	$3, 4 \times 10^{-1}$
rs3740688	11	47380340	*SPI1*[h]	T/G	0.448	0.91 (0.89-0.94)	9.7×10^{-11}	0.93 (0.88-0.97)	1.2×10^{-3}	0.92 (0.89-0.94)	5.4×10^{-13}	$4, 4 \times 10^{-1}$
rs7933202	11	59936926	*MS4A2*	A/C	0.391	0.89 (0.86-0.92)	2.2×10^{-15}	0.90 (0.86-0.95)	1.6×10^{-5}	0.89 (0.87-0.92)	1.9×10^{-19}	$27, 5 \times 10^{-2}$
rs3851179	11	85868640	*PICALM*	C/T	0.356	0.89 (0.86-0.91)	5.8×10^{-16}	0.85 (0.81-0.89)	6.1×10^{-11}	0.88 (0.86-0.90)	6.0×10^{-25}	$0, 8 \times 10^{-1}$
rs11218343	11	121435587	*SORL1*	T/C	0.040	0.81 (0.76-0.88)	2.7×10^{-8}	0.77 (0.68-0.87)	1.8×10^{-5}	0.80 (0.75-0.85)	2.9×10^{-12}	$7, 3 \times 10^{-1}$
rs17125924	14	53391680	*FERMT2*	A/G	0.093	1.13 (1.08-1.19)	6.6×10^{-7}	1.15 (1.06-1.25)	5.0×10^{-4}	1.14 (1.09-1.18)	1.4×10^{-9}	$8, 3 \times 10^{-1}$
rs12881735	14	92932828	*SLC24A4*	T/C	0.221	0.92 (0.88-0.95)	4.9×10^{-7}	0.92 (0.87-0.97)	4.3×10^{-3}	0.92 (0.89-0.94)	7.4×10^{-9}	$0, 6 \times 10^{-1}$
rs3752246	19	1056492	*ABCA7*	C/G	0.182	1.13 (1.09-1.18)	6.6×10^{-10}	1.18 (1.11-1.25)	4.7×10^{-8}	1.15 (1.11-1.18)	3.1×10^{-16}	$0, 5 \times 10^{-1}$
rs429358	19	45411941	*APOE*	T/C	0.216	3.32 (3.20-3.45)	1.2×10^{-881}	*APOE* region not carried forward to replication stage				
rs6024870	20	54997568	*CASS4*	G/A	0.088	0.88 (0.84-0.93)	1.1×10^{-6}	0.90 (0.82-0.97)	9.0×10^{-3}	0.88 (0.85-0.92)	3.5×10^{-8}	$0, 9 \times 10^{-1}$
New genome-wide significant loci reaching significance												
rs7920721	10	11720308	*ECDH3*	A/G	0.389	1.08 (1.05-1.11)	1.9×10^{-7}	1.07 (1.02-1.12)	3.2×10^{-3}	1.08 (1.05-1.11)	2.3×10^{-9}	$0,8 \times 10^{-1}$
rs138190086	17	61538148	*ACE*	G/A	0.020	1.29 (1.15-1.44)	7.5×10^{-6}	1.41 (1.18-1.69)	1.8×10^{-4}	1.32 (1.20-1.45)	7.5×10^{-9}	$0, 9 \times 10^{-1}$
Previous genome-wide significant loci not reaching significance												
rs190982	5	88223420	*MEF2C*	A/G	0.390	0.95 (0.92-0.97)	2.8×10^{-4}	0.93 (0.89-0.98)	2.7×10^{-3}	0.94 (0.92-0.97)	2.8×10^{-6}	$0, 6 \times 10^{-1}$
rs4723711	7	37844263	*NME8*	A/T	0.356	0.95 (0.92-0.98)	2.7×10^{-4}	0.91 (0.87-0.95)	1.0×10^{-4}	0.94 (0.91-0.96)	2.8×10^{-7}	$0, 5 \times 10^{-1}$

[a]Variants showing the best level of association after meta-analysis of stages 1 and 2.

[b]Build 37, assembly hg19.

[c]Based on position of top SNP in reference to the refSeq assembly

[d]Average in the discovery sample.

[e]Calculated with respect to the minor allele.

[f]Cochran's Q test

[g]Previously the ZCWPW1 locus.

[h]Previously the CELF1 locus.

Table 2. Summary of discovery Stage 1, Stage 2, Stage 3 (A and B), and overall meta-analyses results of potential novel loci. Novel loci were defined as loci not reported in Lambert et al. 2013 with 1) a Stage 1+2 Meta $P < 5 \times 10^{-7}$ (9 variants after excluding *TREM2*) (Stage 3A), or 2) a MAF < 0.05 and Stage 1 $P < 1 \times 10^{-5}$ or MAF ≥ 0.05 and Stage 1 $P < 5 \times 10^{-6}$ for genome regions not covered on the Stage 2 custom array (Stage 3B).

Stage 3A						**Stage 1 + 2 (n=82,771)**		**Stage 3A (n=11,666)**		**Overall (n=94,437)**	
SNP[a]	**Chr.**	**Position[b]**	**Closest gene[c]**	**Major/Minor allele**	**MAF[e]**	**OR (95% CI)[f]**	***P***	**OR (95% CI)[f]**	***P***	**OR (95% CI)[f]**	**Meta *P***
rs4735340	8	95976251	*NDUFAF6*	T/A	0.476	0.94 (0.92-0.96)	3.4×10^{-7}	0.92 (0.83-1.02)	9.7×10^{-2}	0.94 (0.92-0.96)	9.2×10^{-8}
rs7920721[g]	10	11720308	*ECHDC3*	A/G	0.390	1.08 (1.05-1.11)	2.30×10^{-9}	1.11 (1.04-1.18)	1.5×10^{-3}	1.08 (1.06-1.11)	1.8×10^{-11}
rs7295246	12	43967677	*ADAMTS20*	T/G	0.413	1.07 (1.04-1.09)	2.7×10^{-7}	1.02 (0.96-1.09)	4.5×10^{-1}	1.06 (1.04-1.08)	3.9×10^{-7}
rs10467994	15	51008687	*SPPL2A*	T/C	0.333	0.97 (0.87-1.08)	3.9×10^{-7}	0.97 (0.87-1.08)	6.2×10^{-1}	0.94 (0.92-0.96)	4.3×10^{-7}
rs593742	15	59045774	*ADAM10*	A/G	0.295	0.93 (0.91-0.96)	1.3×10^{-7}	0.91 (0.85-0.98)	1.5×10^{-2}	0.93 (0.91-0.95)	6.8×10^{-9}
rs7185636	16	19808163	*IQCK*	T/C	0.180	0.92 (0.89-0.95)	8.4×10^{-8}	0.94 (0.86-1.01)	1.1×10^{-1}	0.92 (0.89-0.95)	2.4×10^{-8}
rs2632516	17	56409089	*MIR142/TSPOAP1-AS1*[d]	G/C	0.440	0.94 (0.92-0.96)	2.3×10^{-7}	0.91 (0.82-1.01)	7.5×10^{-2}	0.94 (0.91-0.96)	5.3×10^{-8}
rs138190086	17	61538148	*ACE*	G/A	0.020	1.32 (1.20-1.45)	7.45×10^{-9}	1.17 (0.92-1.48)	2.1×10^{-1}	1.30 (1.19-1.42)	5.3×10^{-9}
rs2830500	21	28156856	*ADAMTS1*	C/A	0.308	0.93 (0.91-0.96)	7.3×10^{-8}	0.95 (0.88-1.02)	1.3×10^{-1}	0.93 (0.91-0.96)	2.6×10^{-8}
Stage 3B						**Stage 1 (n=63,926)**		**Stage 3B (n=30,511)[h]**		**Overall (n=94,437)[h]**	
SNP[a]	**Chr.**	**Position[b]**	**Closest gene[c]**	**Major/Minor allele**	**MAF[e]**	**OR (95% CI)[f]**	***P***	**OR (95% CI)[f]**	***P***	**OR (95% CI)[f]**	**Meta *P***
rs71618613	5	29005985	*SUCLG2P4*	A/C	0.010	0.68 (0.57-0.80)	9.8×10^{-6}	0.76 (0.63-0.93)	6.8×10^{-3}	0.71 (0.63-0.81)	3.3×10^{-7}
rs35868327	5	52665230	*FST*	T/A	0.013	0.69 (0.59-0.80)	7.8×10^{-7}	0.58 (0.29-1.17)	0.126	0.68 (0.59-0.79)	2.6×10^{-7}
rs114812713	6	41034000	*OARD1*	G/C	0.030	1.35 (1.24-1.47)	4.5×10^{-12}	1.23 (1.06-1.42)	7.2×10^{-3}	1.32 (1.22-1.42)	2.1×10^{-13}
rs62039712	16	79355857	*WWOX*	G/A	0.116	1.17 (1.10-1.23)	1.2×10^{-7}	1.14 (0.96-1.36)	0.129	1.16 (1.10-1.23)	3.7×10^{-8}

[a]SNPs showing the best level of association after meta-analysis of stages 1, 2 and 3.
[b]Build 37, assembly hg19.
[c]Based on position of top SNP in reference to the refSeq assembly.
[d]Variant is annotated to both gene features.
[e]Average in the discovery sample.
[f]Calculated with respect to the minor allele.
[g]Recently identified as a LOAD locus in two separate 2017 studies
[h]Sample sizes for these loci are smaller (Overall n=89,769 for *SUCLG2P4*, 65,230 for *LOC257396,FST*, and 69,898 for *WWOX*)

Table 3. Significant pathways (q-value≤0.05) from MAGMA pathway analysis for common SNV and rare SNV subsets.

Pathway	N genes in pathway in dataset	Common SNVs *P**	Common SNVs q-value	Rare SNVs *P**	Rare SNVs q-value	Pathway description
GO:65005	20	1.45E-07*	9.53E-04	6.76E-02	8.42E-01	protein-lipid complex assembly
GO:1902003	10	4.56E-07*	1.49E-03	4.94E-02	8.42E-01	regulation of beta-amyloid formation
GO:32994	39	1.16E-06*	2.54E-03	1.78E-02	8.17E-01	protein-lipid complex
GO:1902991	12	3.54E-06*	5.80E-03	5.66E-02	8.42E-01	regulation of amyloid precursor protein catabolic process
GO:43691	17	5.55E-06*	6.75E-03	3.08E-02	8.17E-01	reverse cholesterol transport
GO:71825	35	6.18E-06*	6.75E-03	1.27E-01	8.42E-01	protein-lipid complex subunit organization
GO:34377	18	1.64E-05*	1.53E-02	1.82E-01	8.42E-01	plasma lipoprotein particle assembly
GO:48156	10	3.19E-05*	2.61E-02	7.77E-01	8.54E-01	tau protein binding
GO:2253	382	6.32E-05*	4.60E-02	2.09E-01	8.42E-01	activation of immune response

*Significant after FDR-correction (q-value≤0.05)

Table 4. Top results of pathway analysis of Aβ-beta centered biological network from Campion et al (see Supplementary Table 12 for full results).

Category	Subcategory	N Genes	Common SNVs *P* 0kb	Common SNVs *P* 35kb-10kb	Rare SNVs *P* 0kb	Rare SNVs *P* 35kb-10kb
Aβ -centered biological network (all genes)	--	331	2.27E-04*	1.54E-04*	8.26E-01	5.19E-01
Clearance and degradation of Aβ	--	74	2.18E-04*	3.27E-03	3.13E-01	5.11E-01
Clearance and degradation of Aβ	Microglia	47	2.24E-04*	1.83E-02	2.49E-01	6.87E-01
Aggregation of Aβ	--	35	7.09E-04*	9.93E-03	9.02E-02	1.68E-01
Aggregation of Aβ	Miscellaneous	21	1.08E-03*	3.38E-02	9.53E-02	1.90E-01
APP processing and trafficking	Clathrin/caveolin-dependent endocytosis	10	1.19E-03	1.15E-02	3.64E-01	1.84E-01
Mediator of Aβ toxicity	--	51	3.82E-02	4.69E-02	5.89E-01	5.70E-01
Mediator of Aβ toxicity	Calcium homeostasis	6	6.90E-02	1.21E-01	3.96E-01	2.54E-01
Mediator of Aβ toxicity	Miscellaneous	3	7.61E-02	2.35E-02	9.79E-01	7.61E-01
Clearance and degradation of Aβ	Enzymatic degradation of Aβ	15	7.77E-02	2.63E-02	6.10E-01	2.95E-01
Mediator of Aβ toxicity	Tau toxicity	20	9.03E-02	3.48E-01	7.17E-01	6.85E-01
Aggregation of Aβ	Chaperone	9	1.52E-01	3.09E-01	1.98E-01	1.13E-02

*Significant after Bonferroni correction for 33 pathway sets tested

Table 5. Top prioritized genes of 400 genes located in genome-wide significant loci. The criteria include: 1) deleterious coding, loss-of-function or splicing variant in gene, 2) significant gene-based test, 3) expression in a tissue relevant to AD (astrocytes, neurons, microglia/macrophages, oligodendrocytes), 4) HuMi microglial-enriched gene, 5) having an eQTL effect on the gene in any tissue, in AD relevant tissue, and/or a co-localized eQTL, 6) being involved in a biological pathway enriched in AD (from the current study), 7) expression correlated with BRAAK stage, and 8) differential expression in 1+ Alzheimer disease (AD) study. Novel genome-wide loci from the current study are listed first, followed by known genome-wide loci. Each category is assigned equal weight of 1, with the priority score equaling the sum of all categories. Colored fields indicate the gene meets the criteria. Genes with a priority score ≥ 4 are listed for each locus. If no gene reached a score of ≥ 5 in a locus, then the top ranked gene(s) is listed.

Evidence Type				Exonic		Tissue Expression		eQTL			Pathway	Clinical Expression	
Locus	Number of Genes in Locus	Prioritized Gene(s)	Priority Score	Coding or Splicing Change	Rare Variant Burden	LOAD Tissue Expression	Microglia-enriched Gene	AD-relevant tissue eQTL	eQTL in any tissue type	Evidence of colocalization	Enriched Pathway	BRAAK Stage Association	DEG Evidence
Novel genome-wide loci													
ADAM10	11	*ADAM10*	5										
IQCK	12	*IQCK*	6										
ACE	22	*PSMC5*	4										
ADAMTS1	3	*ADAMTS1*	4										
WWOX	3	*MAF*	2										
		WWOX	2										
Known genome-wide loci													
CR1	12	*CR1*	7										
		CD55	6										
		YOD1	5										
BIN1	9	*BIN1*	6										
INPP5D	11	*INPP5D*	7										
HLA-DRB1 [†]	46	*HLA-DRB1*	7										
		PSMB8	7										
		C4A	6										
		GPSM3	6										
		HLA-DPA1	6										
		HLA-DQA1	6										
		HLA-DRA	6										
		HLA-DRB5	6										
		PSMB9	6										
TREM2	21	*TREM2*	6										
CD2AP	8	*CD2AP*	5										
NYAP1	53	*AGFG2*	6										
		PILRA	6										
		EPHB4	5										
		C7orf43	5										
		GAL3ST4	5										
		ZKSCAN1	5										
EPHA1	23	*FAM131B*	5										
PTK2B	6	*PTK2B*	5										
CLU	8	*CLU*	6										
ECHDC3	8	*ECHDC3*	4										
SPI1	23	*PSMC3*	6										
		ACP2	5										
		C1QTNF4	5										
		CELF1	5										
		MTCH2	5										
		NDUFS3	5										
		NUP160	5										
		SPI1	5										
MS4A2	24	*MS4A6A*	8										
		MS4A7	6										
		MS4A4A	5										
PICALM	13	*EED*	5										
		PICALM	5										
SORL1	4	*SORL1*	5										
FERMT2	9	*STYX*	5										
SLC24A4	10	*RIN3*	7										
ABCA7	50	*ABCA7*	7										
		HMHA1	6										
		CNN2	5										
		WDR18	5										
CASS4	11	*CASS4*	5										

[†]Genes with rank 6 or above are shown only. An additional 4 genes in *HLA-DRB1* have a priority rank of 5.

Meta-analysis of genetic association with diagnosed Alzheimer's disease identifies novel risk loci and implicates Abeta, Tau, immunity and lipid processing - Methods

Samples. All stage I meta-analysis samples are from four Consortia: the Alzheimer's Disease Genetics Consortium (ADGC), the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium, the European Alzheimer's Disease Initiative (EADI), and the Genetic and Environmental Risk in Alzheimer's Disease (GERAD) Consortium. Summary demographics of all 46 case-control studies from the four consortia are described in **Supplementary Tables 1 and 2**. Written informed consent was obtained from study participants or, for those with substantial cognitive impairment, from a caregiver, legal guardian or other proxy. Study protocols for all cohorts were reviewed and approved by the appropriate institutional review boards. Further details of all cohorts can be found in the **Supplementary Note**.

Pre-imputation genotype chip quality control. Standard quality control (QC) was performed on all datasets individually, including exclusion of individuals with low call rate, individuals with a high degree of relatedness and variants with low call rate. Individuals with non-European ancestry according to principal components (PCs) analysis of ancestry informative markers were excluded from the further analysis.

Imputation and pre-analysis quality control. Following genotype chip QC, each dataset was phased and imputed with data to the 1000 Genomes Project (phase 1 integrated release 3, March 2012)[1] using SHAPEIT/IMPUTE2[2,3] or MaCH/Minimac[4,5] software (**Supplementary Table 3**). All reference population haplotypes were used for the imputation as this method improves accuracy of imputation for low-frequency variants[6]. Common variants (MAF ≥ 0.01%) with an r^2 or an information measure < 0.40 from MaCH and IMPUTE2 were excluded from further analyses. Rare variants (MAF < 0.01%) with a 'global' weighted imputation quality score of < 0.70 were also excluded from analyses. This score was calculated by weighting each variants MACH/IMPUTE2 imputation quality score by study sample size and combining these weighted scores for use as a post-analysis filter. We also required the presence of each variant in 30% of AD cases and 30% of controls across all datasets.

Stage 1 Association Analysis and Meta-analysis. The Stage 1 discovery meta-analysis was followed by Stage 2, and Stage 3 (A and B) replication analyses. Stage 2 was data from a custom array with 11,632 assays selected as variants with $P < 10^{-3}$ from our 2013 work[7]. Genotypes were determined for 8,362 cases and 10,483 controls (**Supplementary Table 4**). Stage 3A was

conducted for variants selected as novel loci from meta-analyses of Stages 1 and 2 with $P < 5 \times 10^{-7}$ (9 variants) and variants that were previously significant ($P < 5 \times 10^{-8}$) that were not genome-wide significant after Stages 1 and 2 (2 variants) (4,930 cases and 6,736 controls) (**Supplementary Table 5**). Stage 3B, which combined samples from Stage 2 and 3A, analysis was conducted for variants with MAF < 0.05 and $P < 1 \times 10^{-5}$ or variants with MAF ≥ 0.05 and $P < 5 \times 10^{-6}$ from genome regions not covered on the Stage 2 custom array (13,292 cases and 17,219 controls) (**Supplementary Table 7**). For Stages 1, 2, and 3, samples did not overlap.

Stage 1 single variant-based association analysis was conducted on genotype dosages modeling for an additive genotype model and adjusting for age (defined as age-at-onset for cases and age-at-last exam for controls), sex and population substructure using PCs[8]. The score test was implemented on all case-control datasets. This test was shown to be optimal for meta-analysis of rare variants due to its balance between power and control of type 1 error[9]. Family datasets were tested using the R package GWAF[10], with generalized estimating equations (GEE) implemented for common variants (MAF ≥ 0.01), and a general linear mixed effects model (GLMM) implemented for rare variants (MAF < 0.01), per internal data showing behavior of test statistics for GEE was fine for common variants but inflated for rare variants, while GLMM controlled this rare variant inflation. Variants with regression coefficient $|\beta| > 5$ or P value equal to 0 or 1 were excluded from further analysis.

Within-study results for Stage 1 were meta-analyzed in METAL[11] using an inverse-variance based model with genomic control. The meta-analysis was split into two separate analyses based on the study sample size, with all studies being included in the analysis of common variants (MAF ≥ 0.01), and only studies with a total sample size of 400 or greater being included in the rare variant (MAF < 0.01) analysis. We also conducted a second meta-analysis in METAL using a sample-size weighted meta-analysis model. Results of this model were compared to the inverse-variance weighted meta-analysis, and results that differed by more than 3 logs on both *P*-values were removed from further analysis. Regression coefficients for rare variants can at times be unstable[12], and this step attempted to control for these problematic variants by using a second method of meta-analysis that may be less sensitive to certain properties of rare variant analysis. In total, 11 variants were removed through this comparison, and most results showed very little difference in *P*-values between the two methods. An additional 106 variants with high heterogeneity between studies (defined as $I^2 > 75$) were removed. Figures for association signals were generated with LocusZoom software[13]. Genome-wide summary statistics are available from The National Institute on Aging Genetics of Alzheimer's Disease (NIAGADS) website

(https://www.niagads.org/). These analyses were conducted by two independent consortia (ADGC and EADI) and then cross-validated.

Stage 1 summary statistics quality control and analysis. Genomic inflation was calculated for lambda in the GenABEL package[14]. In addition, we performed linkage-disequilibrium score (LDSC) regression via LD Hub v1.9.0[15,16] to calculate the LD-score-regression intercept and derive a heritability estimate for the inverse-variance weighted meta-analysis summary statistics. The *APOE* region (Chr19:45,116,911-46,318,605) was removed to calculate the intercept. Removal of the *APOE* region reduced the heritability estimate slightly from 0.071 (s.e. = 0.011) to 0.0637 (s.e. = 0.009).

LDSC was also employed via the LD-Hub web server to obtain genetic correlation estimates (rg)[17] between LOAD and a wide range of other disorders, diseases, and human traits, including 518 UK BioBank traits[18]. UK BioBank is a large long-term study begun in 2006 in the United Kingdom (UK) which is investigating the contributions of genetic predisposition and environmental exposure (i.e. nutrition, lifestyle, medications) to the development of disease. Approximately 500,000 volunteers aged 40 to 69 have been enrolled in the study, with the stated goal of following their health indicators and exposures for 30 years or more after enrollment. While volunteers in the study are generally healthier than the overall UK population[19], it's large size and comprehensive data collection make the study an invaluable resource for researchers looking to interrogate the combined effect of genetics and environmental factors on disease. Prior to analyses in LD-Hub we removed all SNPs with extremely large effect sizes including the MHC (Chr6:26,000,000-34,000,000) and *APOE* region (Chr19:45,116,911-46,318,60) as outliers can overly influence the regression analyses. A total of 1,180,989 variants were used in the correlation analyses. Statistical significance of the genetic correlations was estimated using a 5% Benjamini-Hochberg false discovery rate (FDR)-corrected *P*-values.

GCTA[20] COJO was used to conduct conditional analysis of the Stage 1 summary statistics and using 28,730 unrelated individuals from the ADGC as a reference panel for calculation of linkage disequilibrium (LD). Methods for how the ADGC reference dataset was created are described elsewhere[21,22]. LDLink[23] was used to assess LD, using all 5 CEU populations as the reference for calculations.

Stage 2 and 3 genotyping, quality control, and analysis. Datasets for Stage 2 analysis were obtained from previous genotyping from Lambert et al. 2013[7] using Illumina iSelect technology. The I-select chip has a total of 11,632 single nucleotide variants passing quality control available

for analysis. 1,633 variants were located in the 24 genome-wide loci (defined by the LD-blocks of the sentinel variants; excluding the *APOE* region), with an average of 68 variants per locus. The most well covered loci were the *HLA-DQB1* locus (763 variants), *M24A2* locus (202 variants), and *PICALM* locus (156 variants); the least covered loci were the *MAF* locus (0 variant), *ADAMTS1* locus (4 variants), and the *INPP5D* locus (5 variants). Eleven variants from Stage 3A were genotyped using Taqman technology. Stage 3B included 23 variants included as part of Sequenom MassArray iPLEX panels and 10 additional variants genotyped using Taqman technology.

Per sample quality checks for genetic sex and relatedness were performed in PLINK. Individuals not matching their reported sex or showing a high degree of relatedness (IBD value of 0.98 or greater) were removed from the analysis. A panel of ancestry-informative markers (AIMs), was used to perform PCA analysis with SMARTPCA from EIGENSOFT 4.2 software[24], and individuals with non-European ancestry were excluded. Variant quality control was also performed separately in each country including removal of variants missing in more than 10% of individuals, having a Hardy-Weinberg P value in controls lower than 1×10^{-6}, or a P value for missingness between cases and controls lower than 1×10^{-6}. Please see Lambert et al. for a more detailed description of the QC procedures followed in Stage 2 analysis. After quality control, 18,845 individuals (8,362 cases and 10,483 controls) were available for the stage 2 analysis. The same quality control measures were applied to data for the Stage 3B variants attained from follow-up genotyping.

Selection of variants for Stage 3B follow-up genotyping. In order to prioritize variants for genotyping in Stage 3B, we first selected all MAF < 0.05 variants with $P < 1 \times 10^{-5}$ or MAF ≥ 0.05 variants with $P < 5 \times 10^{-6}$ in novel loci not covered in the iSelect genotyping from Stage 2 of Lambert et al.[7] A total of 180 variants were considered for follow up due to meeting the *P*-value criteria and not being in an IGAP 2013 locus. 88 of these variants were in a region covered in the replication genotyping chip from 2013 and thus were removed from further consideration. 33 loci remained after their removal, with 19 loci having only one prioritized variant, which we selected for genotyping. Remaining variants in 14 regions with multiple prioritized variants were then annotated with GWAVA[25] and CADD[26] scores (using ANNOVAR[27]), Ensembl Variant Effect Predictor (VEP) Consequences (using Ensembl VEP[28]), GWAS4D[29], RegulomeDB[30], and FANTOM5[31] (using NIAGADS GenomicsDB) in order to rank their functional potential. A CADD score > 10, GWAVA score > 0.5, FATHHM > 0.5, RegulomeDB score < 5 and GWAS4D top p-value score were considered 'functional' in the ranking. The top ranked variant for functional

potential for each locus with multiple variants was selected for further genotyping and analysis. Removal of 59 variants in regions with multiple variants left 33 total variants for follow-up genotyping.

Stage 2 and 3 analyses. Per study analysis for Stage 2 and Stage 3 followed the same analysis procedures described for Stage 1, except covariate adjustments per cohort, where all analyses were adjusted on sex and age apart from the Italian, Swedish, and Gr@ACE cohorts, which were also adjusted for PCs. Within-study results for were meta-analyzed in METAL[11] using an inverse-variance based model.

Characterization of gene(s) and non-coding features in associated loci. We determined the basepair (bp) boundaries of the search space for potential gene(s) and non-coding features in each of the 24 associated loci (excluding *APOE*) using the 'proxy search' mechanism in LDLink[23]. LDLink uses 1000 genomes genotypes to calculate LD for a selected population; in our case all five European population were selected (CEU, TSI, FIN, GBR, and IBS). The boundaries for all variants in LD (r2 ≥ 0.5) with the top associated variant from the stage 2 meta-analysis for each region ±500kb of the ends of the LD blocks (as expression quantitative trait loci (eQTL) controlled genes are typically less than 500kb from their controlling variant[32]) were input into the UCSC genome browser's 'Table Browser' for RefSeq[33] and GENCODEv24[34] genes at each associated locus. The average size of the LD blocks was 123kb.

Identification of potentially causal coding or splicing variants. To identify deleterious coding or splicing variants that may represent causal variants for our genome-wide loci we first used SNIPA[35] to identify variants in high LD (defined as r^2>0.7) with the sentinel variants of the 24 genome-wide loci (excluding *APOE*) (N=1,073). The sentinel variants were defined as the variant with the lowest *P* in each genome-wide locus. We then used Ensembl VEP[36] for annotation of the set of sentinel variants and their proxies. We used BLOSUM62[37], SIFT[38], Polyphen-2[39], CADD[26], Condel[40], MPC[41], and Eigen[42] to predict the pathogenicity of protein-altering exonic variants and MaxEntScan to predict the splicing potential of variants. Splicing variants with high splicing potential according to MaxEntScan[43] and protein coding variants predicted to be deleterious by two or more programs were considered to be potentially causal variants for a locus. It should be noted that while we do include rare variants from imputation in our analyses, we may be missing many rare causal variants in this study.

Identification of genes with rare-variant burden via gene-based testing. We used the summary statistics results of a large whole-exome sequencing (WES) study of LOAD, the

Alzheimer's Disease Sequencing Project (ADSP) case-control study (N = 5,740 LOAD cases and 5,096 cognitively normal controls of NHW ancestry) to identify genes within our genome-wide loci that may be contribute to the association signal through rare deleterious coding, splicing or loss-of-function (LOF) variants. The individuals in the ADSP study largely overlap with individuals in the ADGC and CHARGE cohorts included in our Stage 1 meta-analysis. All 445 protein coding genes within our LD defined genome-wide loci were annotated with the gene-based results from this study. Complete details of the analysis can be found in Bis et al. 2018[44]. Briefly, SKAT-O gene-based testing was implemented with seqMeta[45] using multiple models of adjustment (model 0: PC and sequencing center adjusted; model 1: age, sex, PC and sequencing center adjusted; model 2: age, sex, PC, *APOE* and sequencing center adjusted). Only rare (MAF < 0.05), predicted functional and LOF variants were included in the analyses which employed Ensembl VEP consequence categories (high and moderate) and CADD annotation for filtering of variants for inclusion in the SKAT-O analyses. Four annotation models were considered: 1) only rare variants with "HIGH" (splicing or LOF variants) or "MODERATE" (inframe insertions/deletions, missense variants, and predicted protein altering variants) VEP consequences, 2) only rare variants with "HIGH" VEP consequences, 3) only rare variants with CADD Phred scores > 15 (the median value for all possible canonical splice site changes and non-synonymous variants), and 4) only rare variants with CADD Phred scores > 20. The CADD "Phred-scaled" score is obtained from a ranking of all ~8.6 billion variants from the GRCh37/hg19 reference in terms of magnitude and then transforming these ranks to Phred scores, allowing for example a cutoff of the top 0.1% predicted deleterious variants, which is equivalent to our CADD Phred > 20 cutoff. We corrected the results of these models for the 455 genome-wide loci gene list results using a 1% FDR P as a cutoff for significance.

Regulatory variant and eQTL analysis. To identify potential functional risk variants and genes at each associated locus we first annotated a list of prioritized variants from the 24 associated loci (excluding APOE) (N=1,873). This variant list combined variant in LD with the sentinel variants (r2 ≥ 0.5) using INFERNO[46] LD-expansion (N=1,339) and variants with suggestive significance ($P < 10^{-5}$) and LD ($r^2 \geq 0.5$) with the sentinel variants for the 24 associated loci (excluding *APOE*) (N=1,421 variants). We then identified variants with regulatory potential in this set of variants using four programs that incorporate various annotations to identify likely regulatory variants: RegulomeDB[30], HaploReg v4.1[47,48], GWAS4D[29], and the Ensembl Regulatory Build[49]. We used the ChromHMM (Core 15-state model) as "source epigenomes" for the HaploReg analyses. We used immune (Monocytes-CD14+, GM12878 lymphoblastoid, HSMM myoblast) and brain (NH-A astroctyes) for the Ensembl Regulatory Build analyses. We then used the list of 1,873 prioritized

variants to search for genes functionally linked via eQTLs in LOAD relevant tissues including various brain tissue types and blood tissue types, including all immune-related cell types, most specifically myeloid cells (macrophages and monocytes) and B-lymphoid cells, cell types implicated in LOAD and neurodegeneration by a number of recent studies[50–53]. While their specificity may be lower for identifying AD risk eQTLs, we included whole blood cell studies in our AD relevant tissue class due to their high correlation of eQTLs with AD relevant tissues (70% with brain[54]; 51-70% for monocytes and lymphoblastoid cell lines (LCL) respectively[55]) and their large sample sizes which allow for increased discovery power. The eQTL databases and studies searched included: BRAINEAC[56] (12 brain regions), GTEx v7 (48 tissues)[57], BIOSQTL[58], CommonMind Consortium (dorsolateral prefrontal cortex)[59], and xQTLServer[60] (all via FUMA[61]); the NESDA NTR Conditional eQTL Catalog (whole blood)[62]; and Fairfax et al. 2012 (monocytes and B Cells)[63], Gibbs et al. 2010 (frontal cortex, pons)[64], Lappalainen et al. 2013 (LCL)[65], Montgomery et al. 2010 (LCL)[66], MuTHer (Adipose, LCL, skin)[67], and Zeller et al. 2010 (monocytes)[68] (all via exSNP[69]). An additional eQTL overlap search was conducted with INFERNO[46], where 44 GTEx v6 tissues were searched, with prioritization on the INFERNO tissue classes of brain and blood (see **Supplementary Table 13 for sample sizes of each database/study**).

Formal co-localization testing of our summary Stage 1 results was also conducted using 1) COLOC[70] via INFERNO, and 2) Summary Mendelian Randomization (SMR)-Heidi analysis[71]. The approximate bayes factor (ABF), which was used to assess significance in the INFERNO COLOC analysis, is a summary measure that provides an alternative to the P-value for the identification of associations as significant. SMR)-Heidi analysis, which employs a heterogeneity test (HEIDI test) to distinguish pleiotropy or causality (a single genetic variant affecting both gene expression and the trait) from linkage (two distinct genetic variants in LD, one affecting gene expression and one affecting trait), was also employed for co-localization analysis. Genes located less than 1Mb of the GWAS sentinel variants that pass a 5% Benjamini-Hochberg FDR-corrected p-SMR significance threshold and a p-HEIDI > 0.05 threshold were considered significant. The Westra eQTL[72] summary data and Consortium for the Architecture of Gene Expression (CAGE) eQTL summary data was used for analysis. These datasets, conducted in whole blood, are the largest eQTL studies conducted to date (Westra: discovery phase N = 5,311, replication phase N = 2,775; CAGE: N = 2,765), and while there is some overlap in samples between the two datasets, CAGE provides finer coverage. Recent studies have shown significant overlap (50-70%) between brain and blood eQTL's[54]. The ADGC reference panel dataset referenced above for GCTA COJO analysis was used for LD calculations.

Human brain gene expression analyses. We also evaluated gene expression of all candidate genes in the associated loci, defined as all genes within ±500kb of the sentinel variant LD regions ($r^2 \geq 0.5$) (see **Supplementary Table 8** for a complete list of genes searched), using differential AD gene expression results from AlzBase[73], brain tissue expression from the Brain-RNAseq Database (http://www.brainrnaseq.org/[74,75]), and the HuMi_Aged gene set[76], a set of genes preferentially expressed in aged human brain by microglia. This set of genes was established through RNAseq expression analysis of aged human microglial cells from 10 post-mortem brains, and is enriched for AD genes ($P = 4.1 \times 10^{-5}$)[76]. AlzBase includes transcription data from brain and blood from aging, non-dementia, mild cognitive impairment, early stage AD and late stage AD. Please see ALZBase (http://alz.big.ac.cn/alzBase/Document) for a complete list of studies included in the search. Correlation values for the BRAAK stage expression were taken from the Zhang et al. 2013[77] study of 1,647 post-mortem brain tissues from LOAD patients and nondemented subjects.

Pathway Analysis. Pathway analyses were performed with MAGMA[78], which performs SNP-wise gene analysis of summary statistics with correction for LD between variants and genes to test whether sets of genes are jointly associated with a phenotype (i.e. LOAD), compared to other genes across the genome. Adaptive permutation was used to produce an empirical p-value and a FDR-corrected q-value. Gene-sets used in the analyses were from GO[79,80], KEGG[81,82], REACTOME[83,84], BIOCARTA, and MGI[85] pathways. Analyses were restricted to gene sets containing between 10 and 500 genes, a total of 10,861 sets. Variants were restricted to common variants (MAF≥0.01) and rare variants (MAF<0.01) only for each analysis, and separate analyses for each model included and excluded the *APOE* region (Chr19:45,116,911-46,318,605). Analyses were also performed after removal of all genome-wide significant genes. Primary analyses used a 35-kb upstream/10-kb downstream window around each gene in order to potential regulatory variants for each gene, while secondary analyses was run using a 0-kb window[86]. To test for significant correlation between common and rare variant gene results we performed a gene property analysis in MAGMA, regressing the gene-wide association statistics from rare variants on the corresponding statistics from common variants, correcting for LD between variants and genes using the ADGC reference panel. The Aβ-centered network pathway analysis used a curated list of Aβ processing related genes from Campion et al.[87] Thirty-two Aβ–related gene sets and all 335 genes combined (see Campion et al.[87] for details) were run in MAGMA pathway analysis on both common (MAF ≥ 0.01) and rare (MAF < 0.01) variant summary results. The combined dataset of 28,730 unrelated individuals from the ADGC referenced in the GCTA COJO analysis were used as a reference set for LD calculations in these analyses.

Validation of prioritization method. Evaluation of the prioritization of the risk genes in genome-wide loci was done using STRING[88], and Jensen Diseases[89], Jensen Tissues[90], dbGAP gene sets and the ARCHS4[91] resource via the EnrichR[92] tool. We evaluated both the 400 genes set list and a list of 53 genes with priority score ≥ 5 (adding in *APOE* to both lists as the top gene in the *APOE* locus) using the standard settings for both STRING and EnrichR. We use q-value, which is the adjusted p-value using the Benjamini-Hochberg FDR method with a 5% cutoff for correction for multiple hypotheses testing. We also performed 'differentially expressed gene (DEG)' sets analysis via FUMA[61]. These analyses were performed in order to assess whether our 53 prioritized genes are significantly differentially expressed in certain GTEx v7[57] (30 general tissues and 53 specific tissues) or BrainSpan tissues (11 tissue developmental periods with distinct DEG sets ranging from early prenatal to middle adulthood)[93]. FUMA defines DEG sets by calculating a two-sided t-test per tissue versus all remaining tissue types or developmental periods. Genes with a Bonferonni corrected p-value < 0.05 and absolute log fold change ≥ 0.58) are considered DEGs. Input genes were tested against each of the DEG sets using the hypergeometric test. Significant enrichment is defined by Bonferonni corrected P-value ≤ 0.05.

HLA region analysis. Non-familial datasets from ADGC, EADI and GERAD consortiums were used for HLA analysis. After quality control on the imputation quality, a total of 14,776 cases and 23,047 controls were available for analysis (**Supplementary Table 27**). Within ADGC, GenADA, ROSMAP, TARC1, TGEN2, and a subset of UMCWRMSSM datasets were not imputed as Affymetrix genotyping arrays are not supported by the imputation software.

Imputation of HLA alleles. Two-fields resolution HLA alleles were imputed using the R package HIBAG v1.4[94] and the non-Hispanic White (NHW)-specific training set. This software uses specific combinations of variants to predict HLA alleles. Alleles with an imputation posterior probability lower than 0.5 were considered as undetermined as recommended by the developers of the imputation package. *HLA-A*, *HLA-B*, *HLA-C* class I genes and *HLA-DPB1*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DRB1* class II genes were imputed. Individuals with more than two undetermined HLA alleles were excluded.

Statistical analysis. All analyses were performed in R[95]. Associations of HLA alleles with disease were tested using logistic regressions, adjusting for age, sex and PCs as specified above for the SNP association analysis. Only HLA alleles with a frequency higher than 1% were analyzed. Haplotype estimations and association analyses with disease were performed using 'haplo.glm' function from the haplo.stats R package[96] with age, sex and PCs as covariates. Analysis was performed on 2-loci and 3-loci haplotypes of *HLA-DQA1*, *HLA-DQB1* and *HLA-DRB1* genes.

Haplotypes with a frequency below 1% were excluded from the analysis. Considering the high LD in the MHC region, only haplotypes predicted with a posterior probabilities higher than 0.2 were considered for analysis. Meta-analysis p-values were computed using an inverse variance based model as implemented in METAL software[11]. For haplotypes analysis, only individuals with no undetermined HLA alleles and only datasets with more than 100 cases or controls were included. Adjustments on HLA significant variants and HLA alleles were performed by introducing the variant or alleles as covariates in the regression models. Adjusted p-values were computed using the FDR method and the R 'p.adjust' function, and applied to the meta-analysis p-values. FDR threshold was set to 10%.

Data Availability

Stage 1 data (individual level) for the GERAD cohort can be accessed by applying directly to Cardiff University. Stage 1 ADGC data are deposited in a NIAGADS- and NIA/NIH-sanctioned qualified-access data repository. Stage 1 CHARGE data are accessible by applying to dbGaP for all US cohorts and to Erasmus University for Rotterdam data. AGES primary data are not available owing to Icelandic laws. Genome-wide summary statistics for the Stage 1 discovery are available from The National Institute on Aging Genetics of Alzheimer's Disease (NIAGADS) website (https://www.niagads.org/). Stage 2 and stage 3 primary data are available upon request.

URLs:

Brain RNA-seq Database: http://www.brainrnaseq.org/

Enrichr: http://amp.pharm.mssm.edu/Enrichr/

exSNP: http://www.exsnp.org/

NESDA eQTL catalog: https://eqtl.onderzoek.io/index.php?page=info

FUMA: http://fuma.ctglab.nl/

HLA-PheWas Catalog: https://phewascatalog.org/hla

INFERNO: http://inferno.lisanwanglab.org/index.php

LD-Hub: http://ldsc.broadinstitute.org/ldhub/

STRING: https://string-db.org/

Methods References

1. Genomes Project, C. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491,** 56–65 (2012).

2. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5,** e1000529 (2009).

3. Delaneau, O., Marchini, J. & Zagury, J. F. A linear complexity phasing method for thousands of genomes. *Nat Methods* **9,** 179–181 (2012).

4. Li, Y., Willer, C. J., Ding, J., Scheet, P. & Abecasis, G. R. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* **34,** 816–834 (2010).

5. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G. R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* **44,** 955–9 (2012).

6. Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3* **1,** 457–470 (2011).

7. Lambert, J. C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45,** 1452–8 (2013).

8. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38,** 904–909 (2006).

9. Ma, C. *et al.* Recommended Joint and Meta-Analysis Strategies for Case-Control Association Testing of Single Low-Count Variants. *Genet Epidemiol* **37,** 539–550 (2013).

10. Chen, M.-H. H. & Yang, Q. GWAF: an R package for genome-wide association analyses with family data. *Bioinformatics* **26,** 580–581 (2010).

11. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26,** 2190–2191 (2010).

12. Lee, S., Abecasis, G. R., Boehnke, M. & Lin, X. Rare-Variant Association Analysis: Study Designs and Statistical Tests. *Am. J. Hum. Genet.* **95,** 5–23 (2014).

13. Pruim, R. J. *et al.* LocusZoom: regional visualization of genome-wide association scan

results. *Bioinformatics* **26,** 2336–2337 (2010).

14. Aulchenko, Y. S., Ripke, S., Isaacs, A. & van Duijn, C. M. GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23,** 1294–1296 (2007).

15. Zheng, J. *et al.* LD Hub: A centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33,** 051094 (2017).

16. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47,** 291–295 (2015).

17. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47,** 1236–1241 (2015).

18. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med.* **12,** 1–10 (2015).

19. Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants with Those of the General Population. *Am. J. Epidemiol.* **186,** 1026–1034 (2017).

20. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet* **44,** 369–75, S1-3 (2012).

21. Adams, P. M. *et al.* Assessment of the genetic variance of late-onset Alzheimer's disease. *Neurobiol. Aging* **41,** 1–8 (2016).

22. Boehme, K. L. *et al.* ADGC 1000 Genomes combined data workflow. 1–12 (2014). Available at: http://kauwelab.byu.edu/Portals/22/adgc_combined_1000G_12032014.pdf.

23. Machiela, M. J. & Chanock, S. J. LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* **31,** btv402- (2015).

24. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2,** e190 (2006).

25. Ritchie, G. R., Dunham, I., Zeggini, E. & Flicek, P. Functional annotation of noncoding sequence variants. *Nat Methods* **11,** 294–296 (2014).

26. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* **46,** 310–315 (2014).

27. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38,** e164 (2010).

28. McLaren, W. *et al.* Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **26,** 2069–2070 (2010).

29. Huang, D. *et al.* GWAS4D: multidimensional analysis of context-specific regulatory variant for human complex diseases and traits. *Nucleic Acids Res.* **46,** W114–W120 (2018).

30. Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22,** 1790–7 (2012).

31. Andersson, R. *et al.* An atlas of active enhancers across human cell types and tissues. *Nature* **507,** 455–461 (2014).

32. Zhang, X. *et al.* Synthesis of 53 tissue and cell line expression QTL datasets reveals master eQTLs. *BMC Genomics* **15,** 532 (2014).

33. Pruitt, K. D., Tatusova, T., Brown, G. R. & Maglott, D. R. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res* **40,** D130-5 (2012).

34. Harrow, J. *et al.* GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* **22,** 1760–1774 (2012).

35. Arnold, M., Raffler, J., Pfeufer, a., Suhre, K. & Kastenmuller, G. SNiPA: an interactive, genetic variant-centered annotation browser. *Bioinformatics* **31,** 1334–1336 (2014).

36. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *bioRxiv* 042374 (2016). doi:10.1101/042374

37. Cargill, M. *et al.* Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat Genet* **22,** 231–238 (1999).

38. Ng, P. C. & Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* **31,** 3812–3814 (2003).

39. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations.

Nat Methods **7,** 248–249 (2010).

40. Gonzalez-Perez, A. & Lopez-Bigas, N. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *Am J Hum Genet* **88,** 440–449 (2011).

41. Samocha, K. E. *et al.* Regional missense constraint improves variant deleteriousness prediction. *bioRxiv* 148353 (2017). doi:10.1101/148353

42. Ionita-Laza, I., McCallum, K., Xu, B. & Buxbaum, J. D. A spectral approach integrating functional genomic annotations for coding and noncoding variants. *Nat. Genet.* **advance on,** 214–220 (2016).

43. Yeo, G. & Burge, C. B. Maximum Entropy Modeling of Short Sequence Motifs with Applications to RNA Splicing Signals. *J. Comput. Biol.* **11,** 377–394 (2004).

44. Bis, J. C. *et al.* Whole exome sequencing study identifies novel rare and common Alzheimer's-associated variants involved in immune response and transcriptional regulation. *Mol. Psychiatry* Accepted May (2018).

45. seqMeta: an R Package for meta-analyzing region-based tests of rare DNA variants. 1–17 (2014).

46. Amlie-Wolf, A. *et al.* INFERNO – INFERring the molecular mechanisms of NOncoding genetic variants. *bioRxiv* **Oct 30,** (2017).

47. Ward, L. D. & Kellis, M. HaploReg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* **40,** D930-4 (2012).

48. Ward, L. D. & Kellis, M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res.* **44,** gkv1340 (2015).

49. Zerbino, D. R., Wilder, S. P., Johnson, N., Juettemann, T. & Flicek, P. R. The Ensembl Regulatory Build. *Genome Biol.* **16,** 56 (2015).

50. Sims, R. C. *et al.* Novel rare coding variants in PLCG2, ABI3 and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. *Nat. Genet.* 1373–1387 (2017). doi:10.1038/ng.3916

51. Huang, K. *et al.* A common haplotype lowers PU.1 expression in myeloid cells and delays onset of Alzheimer's disease. *Nat. Neurosci.* **20,** 1052–1061 (2017).

52. Thériault, P., ElAli, A. & Rivest, S. The dynamics of monocytes and microglia in Alzheimer's disease. *Alzheimers. Res. Ther.* **7,** 41 (2015).

53. Raj, T. *et al.* Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. *Science* **344,** 519–23 (2014).

54. Qi, T. *et al.* Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *bioRxiv* 1–26 (2018). doi:10.1101/274472

55. Schramm, K. *et al.* Mapping the genetic architecture of gene regulation in whole blood. *PLoS One* **9,** (2014).

56. Ramasamy, A. *et al.* Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat. Neurosci.* **17,** 1418–1428 (2014).

57. Aguet, F. *et al.* Genetic effects on gene expression across human tissues. *Nature* **550,** 204–213 (2017).

58. Zhernakova, D. V. *et al.* Identification of context-dependent expression quantitative trait loci in whole blood. *Nat. Genet.* **49,** 139–145 (2017).

59. Fromer, M. *et al.* Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat. Neurosci.* **19,** 1442–1453 (2016).

60. Ng, B. *et al.* An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. *Nat. Neurosci.* 1–15 (2017). doi:10.1038/nn.4632

61. Watanabe, K., Taskesen, E., Van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8,** 1–10 (2017).

62. Jansen, R. *et al.* Conditional eQTL Analysis Reveals Allelic Heterogeneity of Gene Expression. *Hum. Mol. Genet.* **26,** 1444–1451 (2017).

63. Fairfax, B. P. *et al.* Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet* **44,** 502–510 (2012).

64. Gibbs, J. R. *et al.* Abundant quantitative trait loci exist for DNA methylation and gene expression in Human Brain. *PLoS Genet.* **6,** 29 (2010).

65. Lappalainen, T. *et al.* Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* **501,** 506–11 (2013).

66. Montgomery, S. B. *et al.* Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature* **464,** 773–777 (2010).

67. Grundberg, E. *et al.* Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet* **44,** 1084–1089 (2012).

68. Zeller, T. *et al.* Genetics and beyond - the transcriptome of human monocytes and disease susceptibility. *PLoS One* **5,** (2010).

69. Yu, C.-H., Pal, L. R. & Moult, J. Consensus Genome-Wide Expression Quantitative Trait Loci and Their Relationship with Human Complex Trait Disease. *Omi. A J. Integr. Biol.* **20,** 400–414 (2016).

70. Giambartolomei, C. *et al.* Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genet.* **10,** e1004383 (2014).

71. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48,** 481–489 (2016).

72. Westra, H.-J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* **45,** 1238–1243 (2013).

73. Bai, Z. *et al.* AlzBase: an Integrative Database for Gene Dysregulation in Alzheimer's Disease. *Mol. Neurobiol.* **53,** 310–319 (2016).

74. Zhang, Y. *et al.* An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* **34,** 11929–11947 (2014).

75. Zhang, Y. *et al.* Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. *Neuron* **89,** 37–53 (2016).

76. Olah, M. *et al.* A transcriptomic atlas of aged human microglia. *Nat. Commun.* **9,** 1–8 (2018).

77. Zhang, B. *et al.* Integrated Systems Approach Identifies Genetic Nodes and Networks in Late-Onset Alzheimer's Disease. *Cell* **153,** 707–720 (2013).

78. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-

Set Analysis of GWAS Data. *PLoS Comput. Biol.* **11,** 1–19 (2015).

79. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25,** 25–29 (2000).

80. Blake, J. A. *et al.* Gene ontology consortium: Going forward. *Nucleic Acids Res.* **43,** D1049–D1056 (2015).

81. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44,** D457–D462 (2016).

82. Ogata, H. *et al.* KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **27,** 29–34 (1999).

83. Fabregat, A. *et al.* The reactome pathway knowledgebase. *Nucleic Acids Res.* **44,** D481–D487 (2016).

84. Croft, D. *et al.* Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res* **39,** D691-7 (2011).

85. Eppig, J. T., Blake, J. a, Bult, C. J., Kadin, J. a & Richardson, J. E. The Mouse Genome Database (MGD): facilitating mouse as a model for human biology and disease. *Nucleic Acids Res.* 1–11 (2014). doi:10.1093/nar/gku967

86. O'Dushlaine, C. *et al.* Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat. Neurosci.* **18,** 199–209 (2015).

87. Campion, D., Pottier, C., Nicolas, G., Le Guennec, K. & Rovelet-Lecrux, A. Alzheimer disease: modeling an Aβ-centered biological network. *Mol. Psychiatry* 1–11 (2016). doi:10.1038/mp.2016.38

88. Szklarczyk, D. *et al.* STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43,** D447–D452 (2015).

89. Pletscher-Frankild, S., Pallejà, A., Tsafou, K., Binder, J. X. & Jensen, L. J. DISEASES: Text mining and data integration of disease-gene associations. *Methods* **74,** 83–89 (2015).

90. Santos, A. *et al.* Comprehensive comparison of large-scale tissue expression datasets. *PeerJ* **3,** e1054 (2015).

91. Lachmann, A. *et al.* Massive Mining of Publicly Available RNA-seq Data from Human and Mouse. *Bioarxiv* 1–9 (2017). doi:10.1101/189092

92. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44,** W90–W97 (2016).

93. Miller, J. A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508,** 199–206 (2014).

94. Zheng, X. *et al.* HIBAG - HLA genotype imputation with attribute bagging. *Pharmacogenomics J.* **14,** 192–200 (2014).

95. R Development Core Team. R: a language and environment for statistical computing.

96. Sinwell, J. & Schaid, D. haplo.stats: Statistical Analysis of Traits and Covariates when Linkage Phase is Ambiguous. (2016).